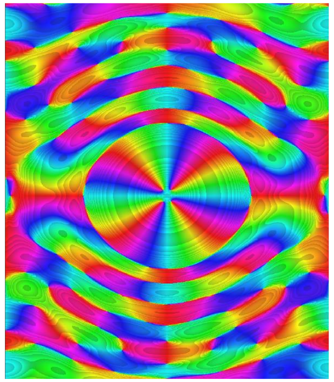
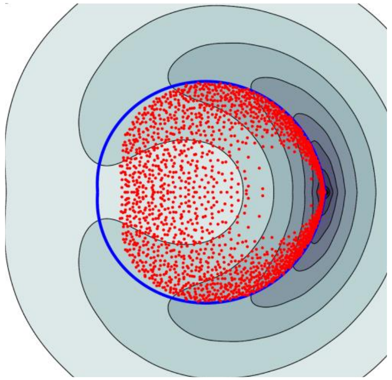
Residual Dynamic Mode Decomposition
Robust and verified Koopmanism!

Matthew Colbrook (m.colbrook@damtp.cam.ac.uk)
University of Cambridge

Joint work with
Lorna Ayton (Cambridge), **Máté Szőke** (Virginia Tech), **Alex Townsend** (Cornell)





Maths

C., Townsend, “*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,*” preprint.
C., Ayton, Szőke, “*Residual Dynamic Mode Decomposition,*” preprint.

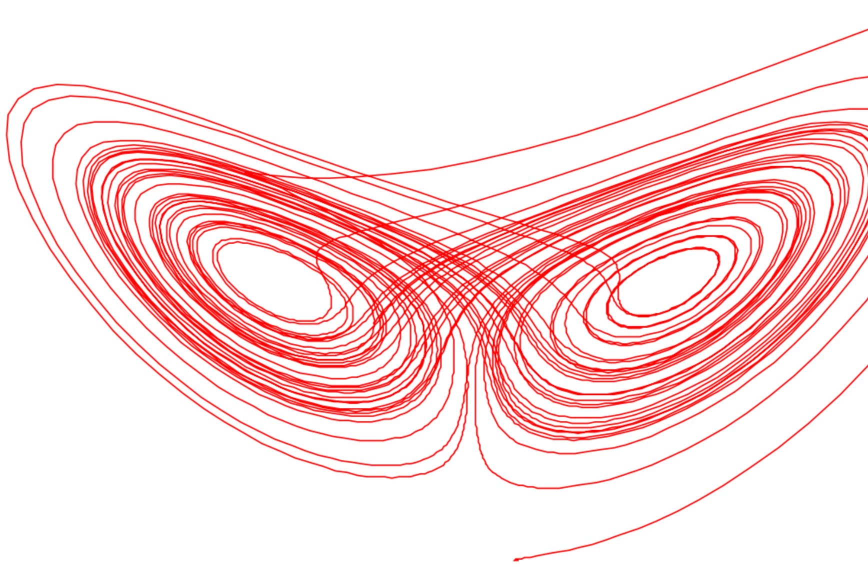
Applications

Data-driven dynamical systems

- State $x \in \Omega \subseteq \mathbb{R}^d$, ***unknown*** function $F: \Omega \to \Omega$ governs dynamics

$$x_{n+1} = F(x_n)$$

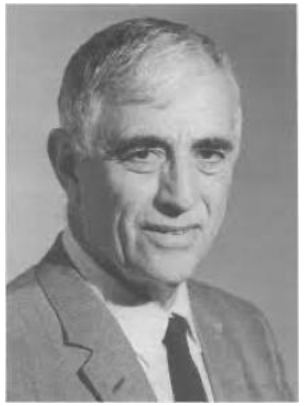
- **Goal:** Learn about system from data $\left\{x^{(m)}, y^{(m)} = F(x^{(m)})\right\}_{m=1}^{M}$
  - **Data:** experimental measurements or numerical simulations
  - E.g., **used for** forecasting, control, design, understanding
- **Applications:** chemistry, climatology, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, plasmas, robotics, video processing, etc.



Poincaré

Operator viewpoint

Koopman

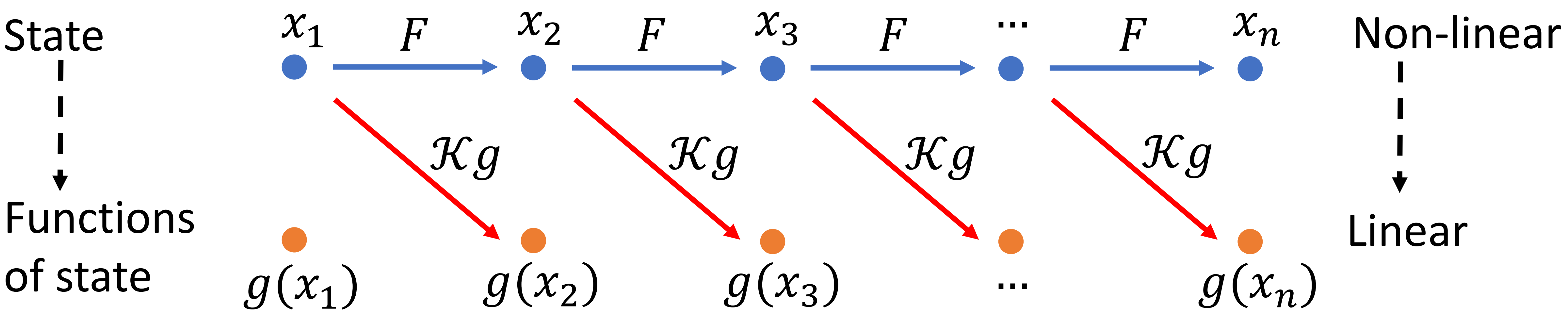


von Neumann

- **Koopman operator** $\mathcal{K}$ acts on <u>functions</u> $g: \Omega \to \mathbb{C}$

$$[\mathcal{K}g](x_n) = g\big(F(x_n)\big) = g(x_{n+1})$$

- $\mathcal{K}$ is ***linear*** but acts on an ***infinite-dimensional*** space.



- Work in $L^2(\Omega, \omega)$ for positive measure ω, with inner product $\langle \cdot, \cdot \rangle$.

- Koopman, "*Hamiltonian systems and transformation in Hilbert space,*" **Proc. Natl. Acad. Sci. USA**, 1931.
- Koopman, v. Neumann, "*Dynamical systems of continuous spectra,*" **Proc. Natl. Acad. Sci. USA**, 1932.

Why is linear (<u>much</u>) easier?

$$\boxed{x_{n+1} = F(x_n)}$$

Long-time dynamics become trivial!

- Suppose $F(x) = Ax, A \in \mathbb{R}^{d\times d}$, $A = V\Lambda V^{-1}$.
- Set $\xi = V^{-1}x$,

$$\xi_n = V^{-1}x_n = V^{-1}A^n x_0 = \Lambda^n V^{-1} x_0 = \Lambda^n \xi_0$$

- Let $w^{\mathrm{T}}A = \lambda w$, set $\varphi(x) = w^{\mathrm{T}}x$,

$$[\mathcal{K}\varphi](x) = w^{\mathrm{T}}Ax = \lambda\varphi(x)$$

Eigenfunction

Much more general (**non-linear** and even **chaotic** F).

Koopman mode decomposition

generalized eigenfunction of $\mathcal{K}$

eigenfunction of $\mathcal{K}$

$$[\mathcal{K}g](x) = g(F(x))$$

$$g(x) = \sum_{\text{eigs } \lambda_j} c_{\lambda_j} \varphi_{\lambda_j}(x) + \int_{[-\pi,\pi]_{\text{per}}} \phi_{\theta,g}(x)\, \mathrm{d}\theta$$

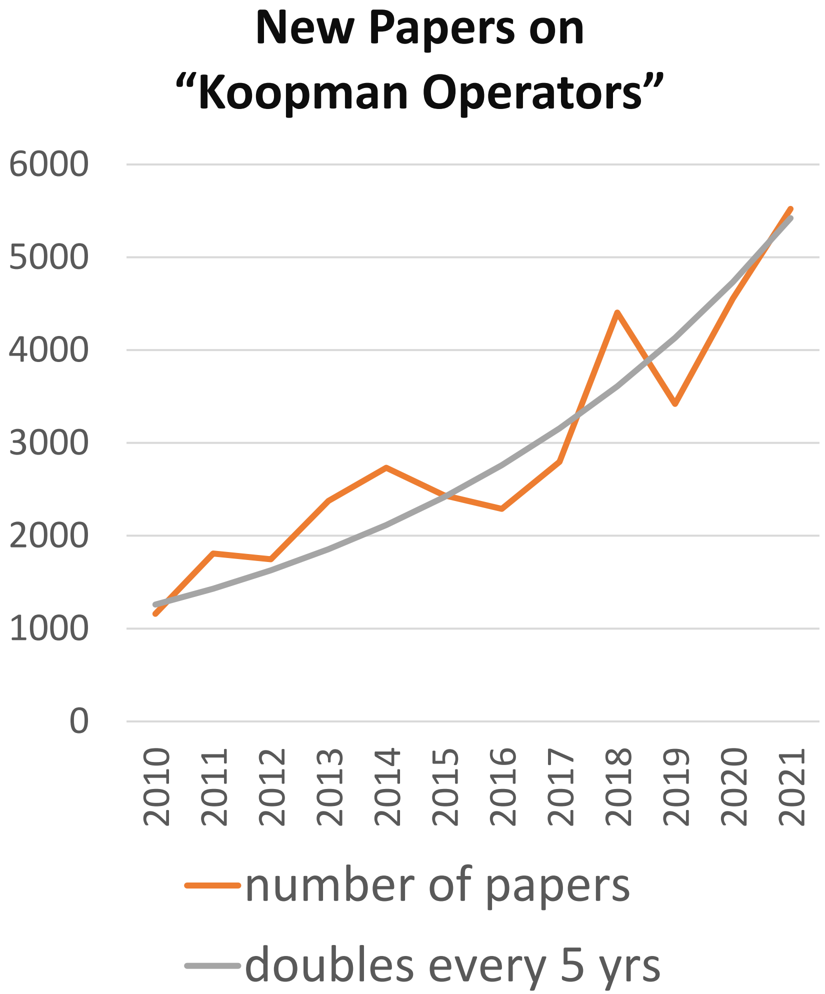
$$g(x_n) = [\mathcal{K}^n g](x_0) = \sum_{\text{eigs } \lambda_j} c_{\lambda_j} {\lambda_j}^n \varphi_{\lambda_j}(x_0) + \int_{[-\pi,\pi]_{\text{per}}} e^{in\theta} \phi_{\theta,g}(x_0)\, \mathrm{d}\theta$$

Encodes: geometric features, invariant measures, transient behavior, long-time behavior, coherent structures, quasiperiodicity, etc.

GOAL: Data-driven approximation of $\mathcal{K}$ and its spectral properties.

- Mezić, "*Spectral properties of dynamical systems, model reduction and decompositions,*" **Nonlinear Dynam.**, 2005.

Koopmania*: A revolution in the big data era?



$\approx$35,000 papers over last decade!

BUT: Computing spectra in infinite dimensions is notoriously hard!

**Wikipedia*: "*its wild surge in popularity is sometimes jokingly called 'Koopmania'*"

Challenges of computing $\mathrm{Spec}(\mathcal{K}) = \{\lambda \in \mathbb{C} : \mathcal{K} - \lambda I \text{ is not invertible}\}$

Truncate: $\mathcal{K} \longrightarrow \mathbb{K} \in \mathbb{C}^{N_K \times N_K}$

1) **"Too much":** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$
2) **"Too little":** Miss parts of $\mathrm{Spec}(\mathcal{K})$
3) **Continuous spectra.**

Verification: Is it right?

Build the matrix: Dynamic Mode Decomposition (DMD)

$$\boxed{\{x^{(m)}, y^{(m)} = F(x^{(m)})\}_{m=1}^{M}}$$

Given dictionary $\{\psi_1, \ldots, \psi_{N_K}\}$ of functions $\psi_j : \Omega \to \mathbb{C}$,

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \psi_k(x^{(m)}) = \left[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \cdots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \cdots & \psi_{N_K}(x^{(M)}) \end{pmatrix}^*}_{\Psi_X} \underbrace{\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix}}_{W} \underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \cdots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \cdots & \psi_{N_K}(x^{(M)}) \end{pmatrix}}_{\Psi_X} \right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = \left[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \cdots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \cdots & \psi_{N_K}(x^{(M)}) \end{pmatrix}^*}_{\Psi_X} \underbrace{\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix}}_{W} \underbrace{\begin{pmatrix} \psi_1(y^{(1)}) & \cdots & \psi_{N_K}(y^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(y^{(M)}) & \cdots & \psi_{N_K}(y^{(M)}) \end{pmatrix}}_{\Psi_Y} \right]_{jk}$$

$$\mathcal{K} \longrightarrow \mathbb{K} = (\Psi_X{}^* W \Psi_X)^{-1} \Psi_X{}^* W \Psi_Y \in \mathbb{C}^{N_K \times N_K}$$

Recall open problems: too much, too little, continuous spectra, verification

- Schmid, *"Dynamic mode decomposition of numerical and experimental data,"* **J. Fluid Mech.**, 2010.
- Rowley, Mezić, Bagheri, Schlatter, Henningson, *"Spectral analysis of nonlinear flows,"* **J. Fluid Mech.**, 2009.
- Kutz, Brunton, Brunton, Proctor, *"Dynamic mode decomposition: data-driven modeling of complex systems,"* **SIAM**, 2016.
- Williams, Kevrekidis, Rowley *"A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,"* **J. Nonlinear Sci.**, 2015.

Residual DMD (ResDMD): Approx. $\mathcal{K}$ and $\mathcal{K}^*\mathcal{K}$

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})}\, \psi_k(x^{(m)}) = [\underbrace{\Psi_X{}^* W \Psi_X}_{G}]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})}\, \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = [\underbrace{\Psi_X{}^* W \Psi_Y}_{K_1}]_{jk}$$

$$\langle \mathcal{K}\psi_k, \mathcal{K}\psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(y^{(m)})}\, \psi_k(y^{(m)}) = [\underbrace{\Psi_Y{}^* W \Psi_Y}_{K_2}]_{jk}$$

Residuals: $g = \sum_{j=1}^{N_K} \mathbf{g}_j \psi_j$, $\|\mathcal{K}g - \lambda g\|^2 \approx \mathbf{g}^*[K_2 - \lambda K_1{}^* - \bar{\lambda} K_1 + |\lambda|^2 G]\mathbf{g}$

- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.
- C., Ayton, Szőke, *"Residual Dynamic Mode Decomposition,"* **J. Fluid Mech.,** under minor rev.
- Code*:* https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition

Quadrature with trajectory data

$$\text{E.g.,}\quad \langle \mathcal{K}\psi_k, \psi_j \rangle = \lim_{M\to\infty} \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})}\, \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})}$$

Three examples:

- **High-order quadrature:** $\{x^{(m)}, w_m\}_{m=1}^{M}$ M-point quadrature rule.
  Rapid convergence. Requires free choice of $\{x^{(m)}\}_{m=1}^{M}$ and small d.
- **Random sampling:** $\{x^{(m)}\}_{m=1}^{M}$ selected at random.
  Large d. Slow Monte Carlo $O(M^{-1/2})$ rate of convergence.
- **Ergodic sampling:** $x^{(m+1)} = F(x^{(m)})$.
  Single trajectory, large d. Requires ergodicity, convergence can be slow.

Most common

ResDMD: avoiding "too much"

$$\mathrm{res}(\lambda, \mathbf{g})^2 = \frac{\mathbf{g}^*\left[K_2 - \lambda {K_1}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}}{\mathbf{g}^* G \mathbf{g}}$$

eigenvectors

eigenvalues

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = G V \Lambda$.
2. For each eigenpair $(\lambda, \mathbf{v})$, compute $\mathrm{res}(\lambda, \mathbf{v})$.
3. **Output:** subset of e-vectors $V_{(\varepsilon)}$ & e-vals $\Lambda_{(\varepsilon)}$ with $\mathrm{res}(\lambda, \mathbf{v}) \leq \varepsilon$ ($\varepsilon =$ input tol).

Theorem (no spectral pollution): Suppose quad. rule converges. Then

$$\limsup_{M \to \infty} \max_{\lambda \in \Lambda_{(\varepsilon)}} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \varepsilon$$

• C., Townsend, "*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,*" preprint.

ResDMD: avoiding “too much”

$$\mathrm{res}(\lambda,\mathbf{g})^2 = \frac{\mathbf{g}^*\left[K_2 - \lambda {K_1}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}}{\mathbf{g}^* G \mathbf{g}}$$

eigenvectors

eigenvalues

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = G V \Lambda$.
2. For each eigenpair $(\lambda, \mathbf{v})$, compute $\mathrm{res}(\lambda, \mathbf{v})$.
3. **Output:** subset of e-vectors $V_{(\varepsilon)}$ & e-vals $\Lambda_{(\varepsilon)}$ with $\mathrm{res}(\lambda, \mathbf{v}) \leq \varepsilon$ ($\varepsilon =$ input tol).

Theorem (no spectral pollution): Suppose quad. rule converges. Then

$$\limsup_{M\to\infty} \max_{\lambda\in\Lambda_{(\varepsilon)}} \|(\mathcal{K}-\lambda)^{-1}\|^{-1} \leq \varepsilon$$

BUT: Typically, does not capture all of spectrum! (“too little”)

• C., Townsend, *“Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,”* preprint.

ResDMD: avoiding "too little"

$$\mathrm{Spec}_\varepsilon(\mathcal{K}) = \bigcup_{\|\mathcal{B}\|\leq\varepsilon} \mathrm{Spec}(\mathcal{K}+\mathcal{B}), \qquad \lim_{\varepsilon\downarrow 0} \mathrm{Spec}_\varepsilon(\mathcal{K}) = \mathrm{Spec}(\mathcal{K})$$

Algorithm 2:

First convergent method for general $\mathcal{K}$

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K\times N_K}$.
2. For z_k in comp. grid, compute $\tau_k = \min_{g=\sum_{j=1}^{N_K}\mathbf{g}_j\psi_j} \mathrm{res}(z_k, g)$, corresponding g_k (gen. SVD).
3. **Output:** $\{z_k : \tau_k < \varepsilon\}$ (approx. of $\mathrm{Spec}_\varepsilon(\mathcal{K})$), $\{g_k : \tau_k < \varepsilon\}$ (ε-pseudo-eigenfunctions).

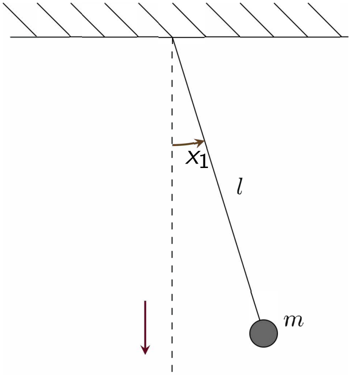
Theorem (full convergence): Suppose the quadrature rule converges.

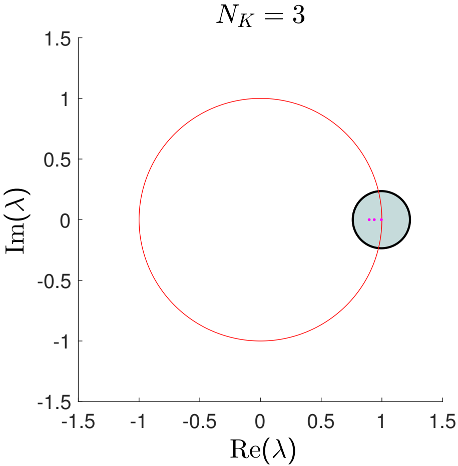
- **Error control:** $\{z_k : \tau_k < \varepsilon\} \subseteq \mathrm{Spec}_\varepsilon(\mathcal{K})$ (as $M\to\infty$)
- **Convergence:** Converges locally uniformly to $\mathrm{Spec}_\varepsilon(\mathcal{K})$ (as $N_K\to\infty$)

• C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.

Example: non-linear pendulum

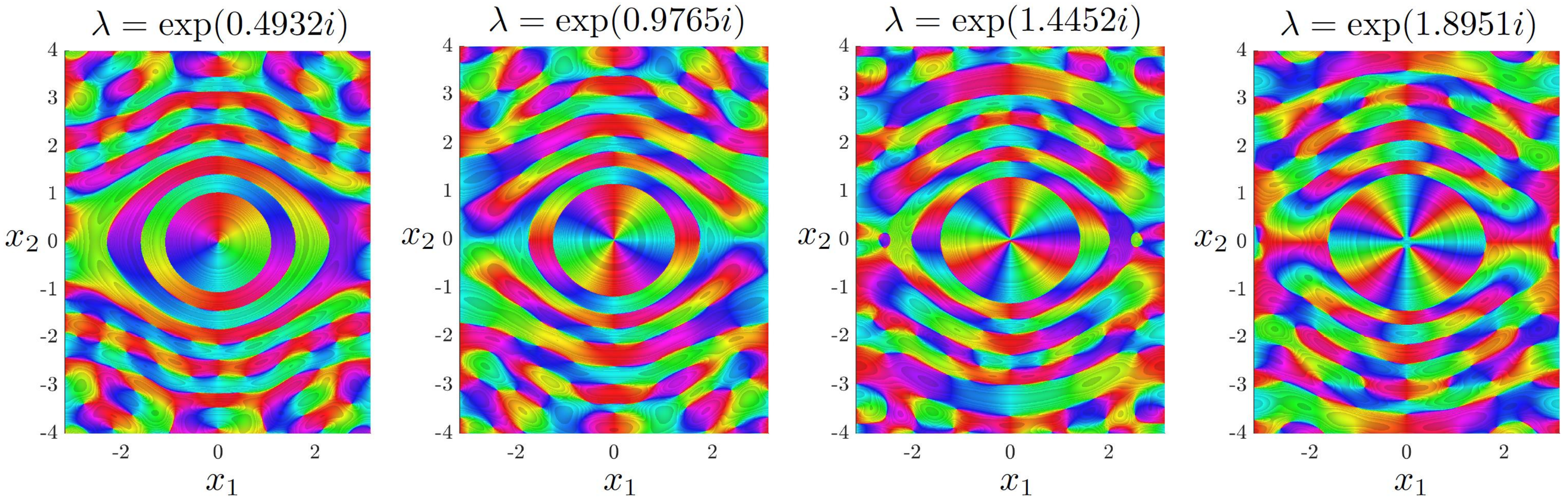
$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -\sin(x_1), \qquad \Omega = [-\pi, \pi]_{\text{per}} \times \mathbb{R}$$





Computed pseudospectra ($\varepsilon = 0.25$). Eigenvalues of $\mathbb{K}$ shown as dots (spectral pollution).

Approximate eigenfunctions



Colour represents complex argument, constant modulus shown as shadowed steps.
All residuals smaller than $\varepsilon = 0.05$ (made smaller by increasing N_K).

The Challenges

~~**1) "Too much":** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$~~ ✓

~~**2) "Too little":** Miss parts of $\mathrm{Spec}(\mathcal{K})$~~ ✓

3) Continuous spectra.

Verification: Is it right?

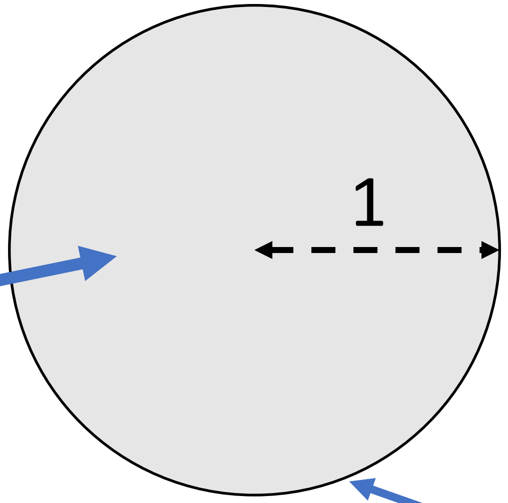
Setup for continuous spectra

No such assumption
was made in first part of talk!

Suppose system is measure-preserving (e.g., Hamiltonian, ergodic, post-transient etc.)

$$\Leftrightarrow \mathcal{K}^*\mathcal{K} = I \text{ (isometry)}$$

$$\Rightarrow \mathrm{Spec}(\mathcal{K}) \subseteq \{z : |z| \leq 1\}$$



spectral
measure
supp. on
boundary

(NB: we consider unitary extensions via Wold decomposition.)

Spectral decomposition of operators

$A \in \mathbb{C}^{n\times n}$ normal $\implies$ O.N. basis of eigenvectors $v_1, \dots, v_n$:

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \qquad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Projector onto Span(v_k)

eigenvalues

Spectral decomposition of operators

$A \in \mathbb{C}^{n\times n}$ normal $\Longrightarrow$ O.N. basis of eigenvectors $v_1, \dots, v_n$:

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \qquad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Projector onto Span(v_k)

eigenvalues

Energy of "v" in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v\rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v.

Spectral decomposition of operators

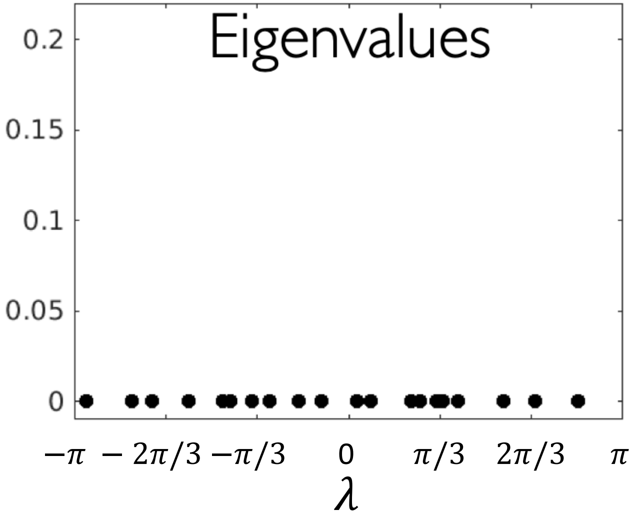
$A \in \mathbb{C}^{n\times n}$ normal $\quad \Longrightarrow \quad$ O.N. basis of eigenvectors $v_1, \dots, v_n$:

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \qquad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Projector onto Span(v_k) $\qquad$ eigenvalues

Energy of "v" in each eigenvector: $\quad \mu_v(\lambda_j) = \langle v_j v_j^* v, v\rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v.



Spectral decomposition of operators

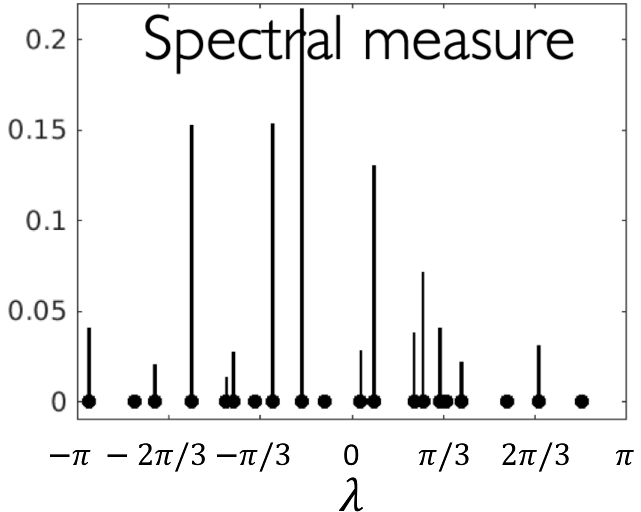
$A \in \mathbb{C}^{n\times n}$ normal $\Longrightarrow$ O.N. basis of eigenvectors $v_1, \dots, v_n$:

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \qquad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Projector onto Span(v_k) eigenvalues

Energy of "v" in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v\rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v.



Spectral decomposition of operators

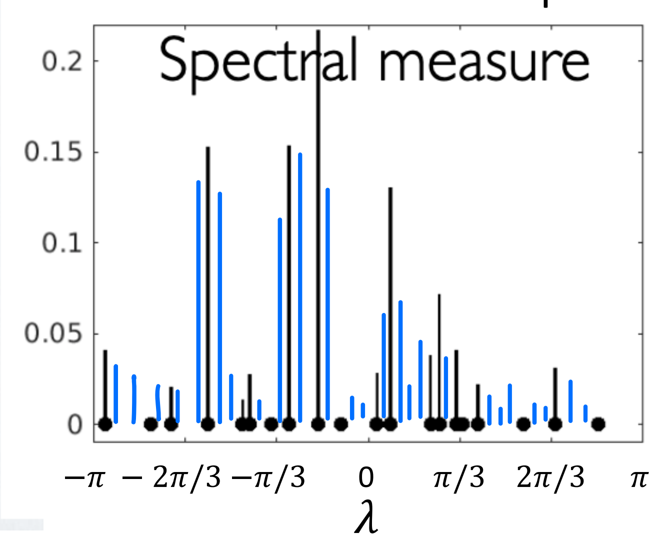
$A \in \mathbb{C}^{n\times n}$ normal $\Longrightarrow$ O.N. basis of eigenvectors $v_1, \ldots, v_n$:

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \qquad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Projector onto $\mathrm{Span}(v_k)$ eigenvalues

Energy of "v" in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v\rangle = \left|v_j^* v\right|^2$

This is called the spectral measure with respect to a vector v.



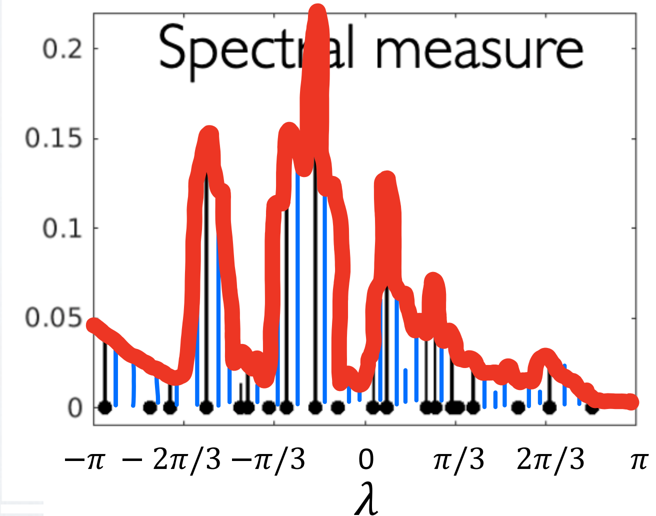
Spectral decomposition of operators

$A \in \mathbb{C}^{n\times n}$ normal $\quad\Longrightarrow\quad$ O.N. basis of eigenvectors $v_1, \ldots, v_n$:

$$v = \left(\sum_{k=1}^{n} \underbrace{v_k v_k^*}_{\text{Projector onto Span}(v_k)}\right) v, \qquad Av = \left(\sum_{k=1}^{n} \underbrace{\lambda_k}_{\text{eigenvalues}} v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v\rangle = \left|v_j^* v\right|^2$

This is called the spectral measure with respect to a vector v.



Spectral decomposition of operators

$A \in \mathbb{C}^{n\times n}$ normal $\Longrightarrow$ O.N. basis of eigenvectors $v_1, \dots, v_n$:

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \qquad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \qquad v \in \mathbb{C}^n$$

Projector onto Span(v_k) eigenvalues

Energy of "v" in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v\rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v.

$\mathcal{K}$ is unitary $\Longrightarrow$ projection-valued measure ξ

$$g = \left(\int_{\mathbb{T}} d\xi(y)\right) g, \qquad \mathcal{K}g = \left(\int_{\mathbb{T}} y\, d\xi(y)\right) g$$

Spectral measure $\quad \nu_g(B) = \langle \xi(B)g, g\rangle$

Spectral decomposition of operators

$A \in \mathbb{C}^{n\times n}$ normal $\Longrightarrow$ O.

$v,$ A

$\mathrm{an}(v_k)$

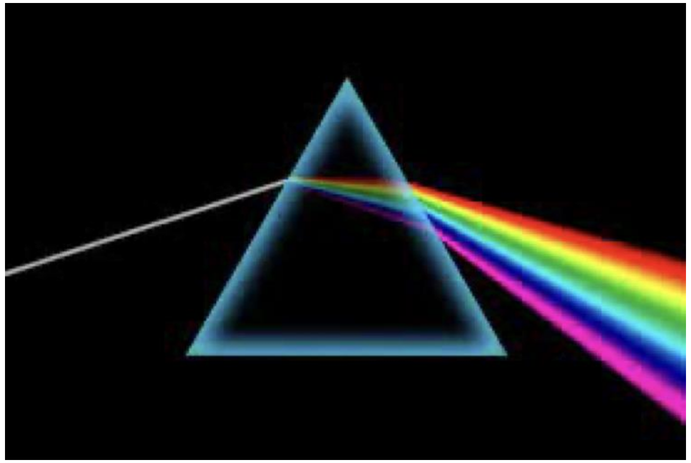
eigenvector:

This is ca

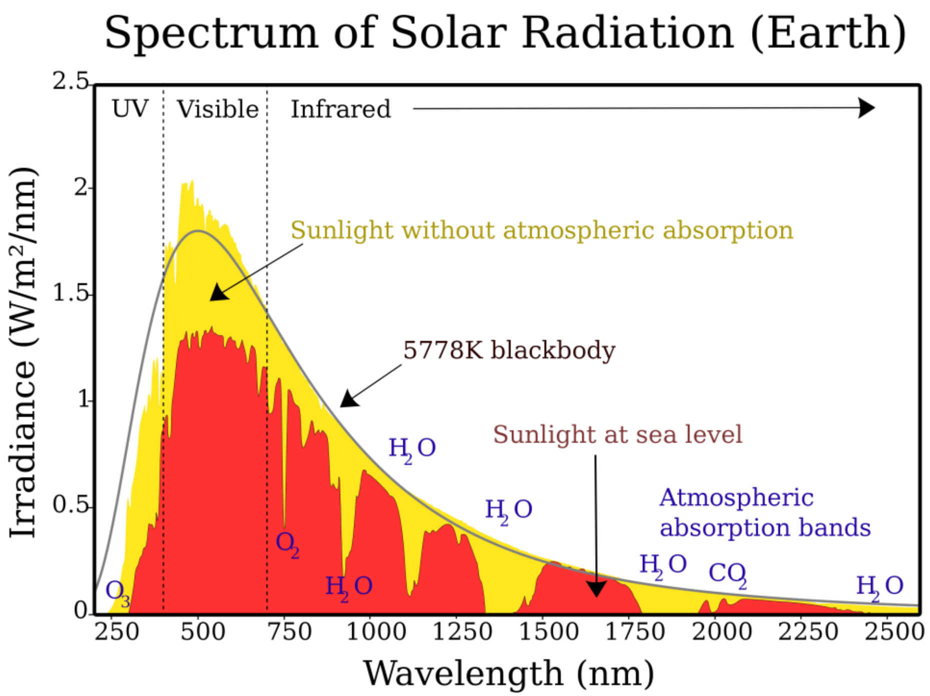
$\Longrightarrow$

$$g = \left(\int_{\mathbb{T}} d\xi(y)\right) g,$$

White light contains a continuous spectrum

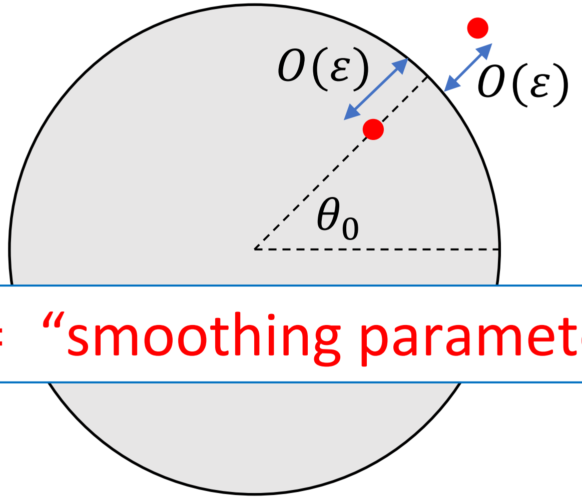


Often interesting to look at
the intensity of each wavelength



Spectral measure $\nu_g(B) = \langle \xi(B)g, g\rangle$

Evaluating spectral measure



$\varepsilon =$ "smoothing parameter"

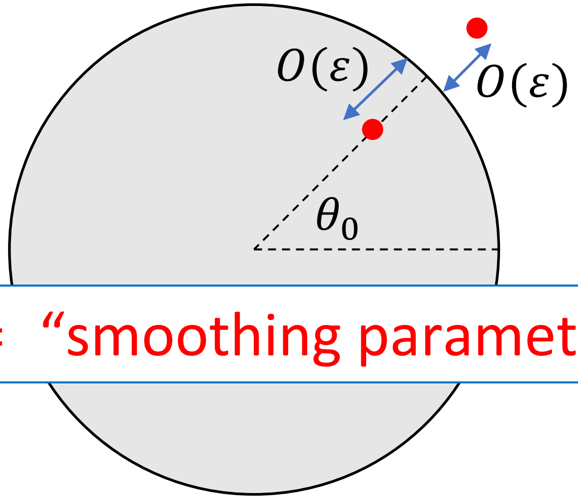
Smoothing convolution

$$[P_\varepsilon * \nu_g](\theta_0) = \int_{[-\pi,\pi]_{\mathrm{per}}} P_\varepsilon(\theta_0 - \theta)\, \mathrm{d}\nu_g(\theta)$$

Poisson kernel for unit disk

$$P_\varepsilon(\theta_0) = \frac{1}{2\pi} \frac{(1+\varepsilon)^2 - 1}{1 + (1+\varepsilon)^2 - 2(1+\varepsilon)\cos(\theta_0)}$$

Evaluating spectral measur



$\varepsilon =$ "smoothing parameter"

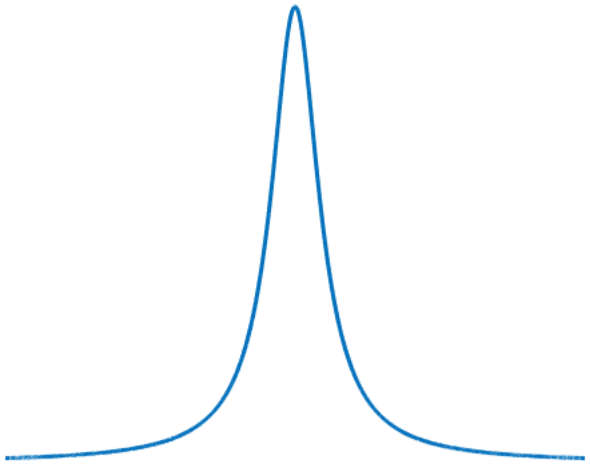
Smoothing cc

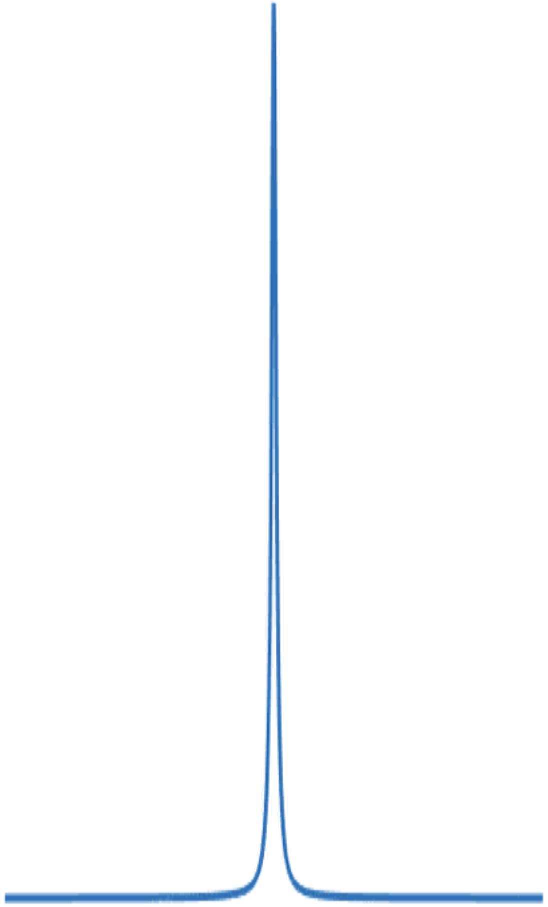
$$[P_\varepsilon * \nu_g](\theta_0) = \int P_\varepsilon(\theta_0 - \theta$$

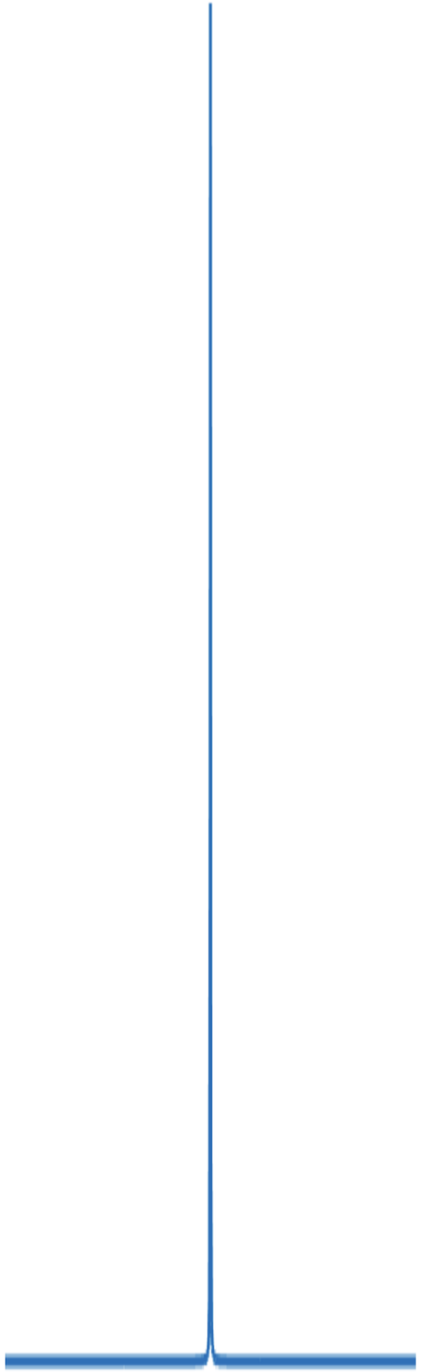
Poisso
unit di

$+ (1 + \varepsilon$

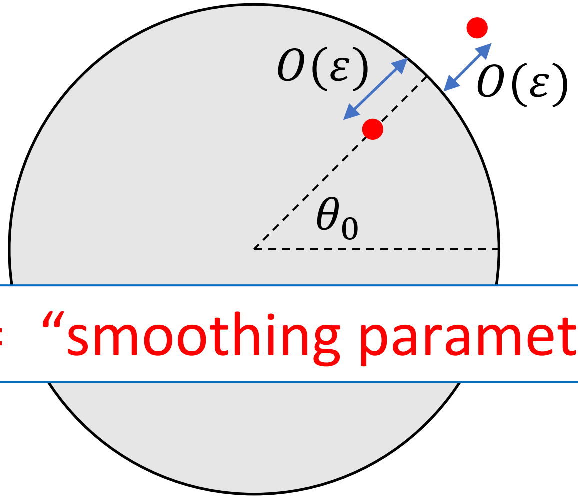
$_0)$







Evaluating spectral measure



$\varepsilon =$ "smoothing parameter"

Smoothing convolution

$$[P_\varepsilon * \nu_g](\theta_0) = \int_{[-\pi,\pi]_{\mathrm{per}}} P_\varepsilon(\theta_0 - \theta)\, \mathrm{d}\nu_g(\theta)$$

Poisson kernel for unit disk

$$P_\varepsilon(\theta_0) = \frac{1}{2\pi} \frac{(1+\varepsilon)^2 - 1}{1 + (1+\varepsilon)^2 - 2(1+\varepsilon)\cos(\theta_0)}$$

$$[P_\varepsilon * \nu_g](\theta_0) = \mathcal{C}_g\left(e^{i\theta_0}(1+\varepsilon)^{-1}\right) - \mathcal{C}_g\left(e^{i\theta_0}(1+\varepsilon)\right)$$

$$\mathcal{C}_g(z) = \int_{[-\pi,\pi]_{\mathrm{per}}} \frac{e^{i\theta}\,\mathrm{d}\nu_g(\theta)}{e^{i\theta} - z} = \begin{cases} \langle (\mathcal{K} - zI)^{-1} g, \mathcal{K}^* g \rangle, & \text{if } |z| > 1 \\ -z^{-1} \langle g, (\mathcal{K} - \bar{z}^{-1} I)^{-1} g \rangle, & \text{if } 0 < |z| < 1 \end{cases}$$

ResDMD computes with error control

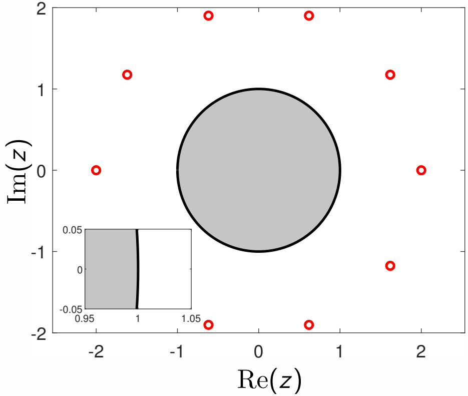
Example

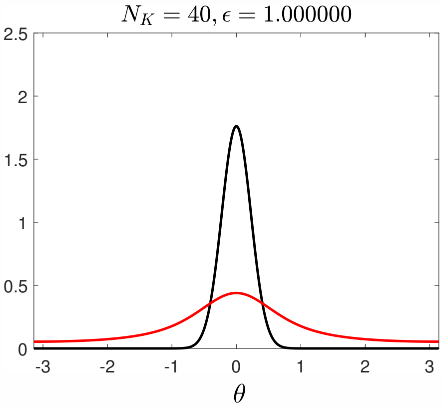
$$\mathcal{K} = \begin{pmatrix} \overline{\alpha_0} & \overline{\alpha_1}\rho_0 & \rho_0\rho_1 & & & \\ \rho_0 & -\overline{\alpha_1}\alpha_0 & -\alpha_0\rho_1 & & & \\ & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & \\ & \rho_2\rho_1 & -\alpha_1\rho_2 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & \ddots \\ & & & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \ddots \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

$$\alpha_j = (-1)^j 0.95^{(j+1)/2}, \qquad \rho_j = \sqrt{1 - |\alpha_j|^2}$$

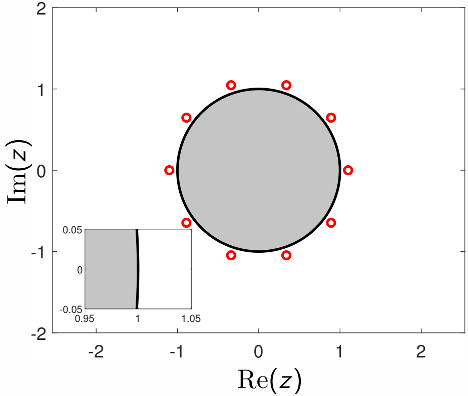
Generalized shift, typical building block of many dynamical systems.

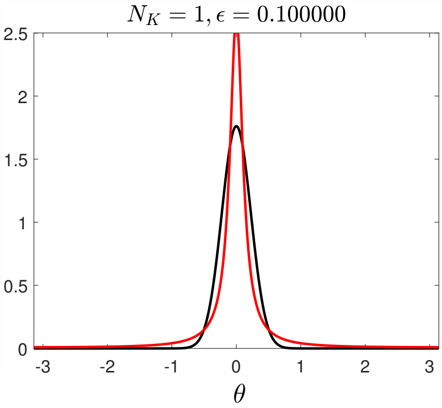
Fix N_K, vary ε: unstable!



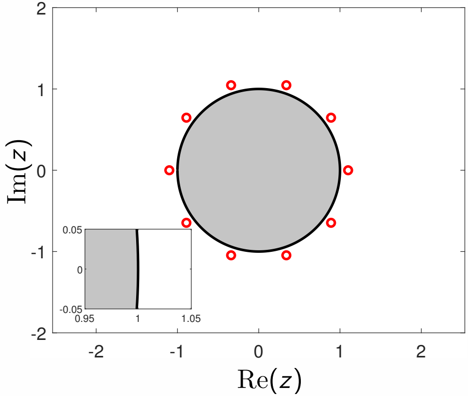


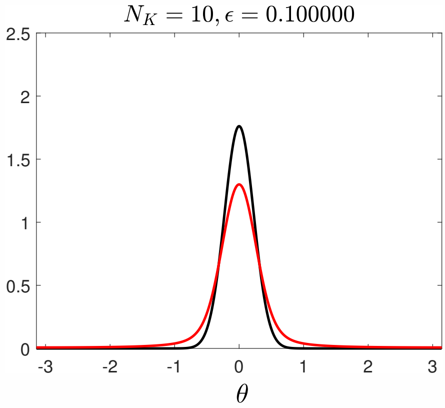
Fix ε, vary N_K: too smooth!





Adaptive: new matrix to compute residuals crucial

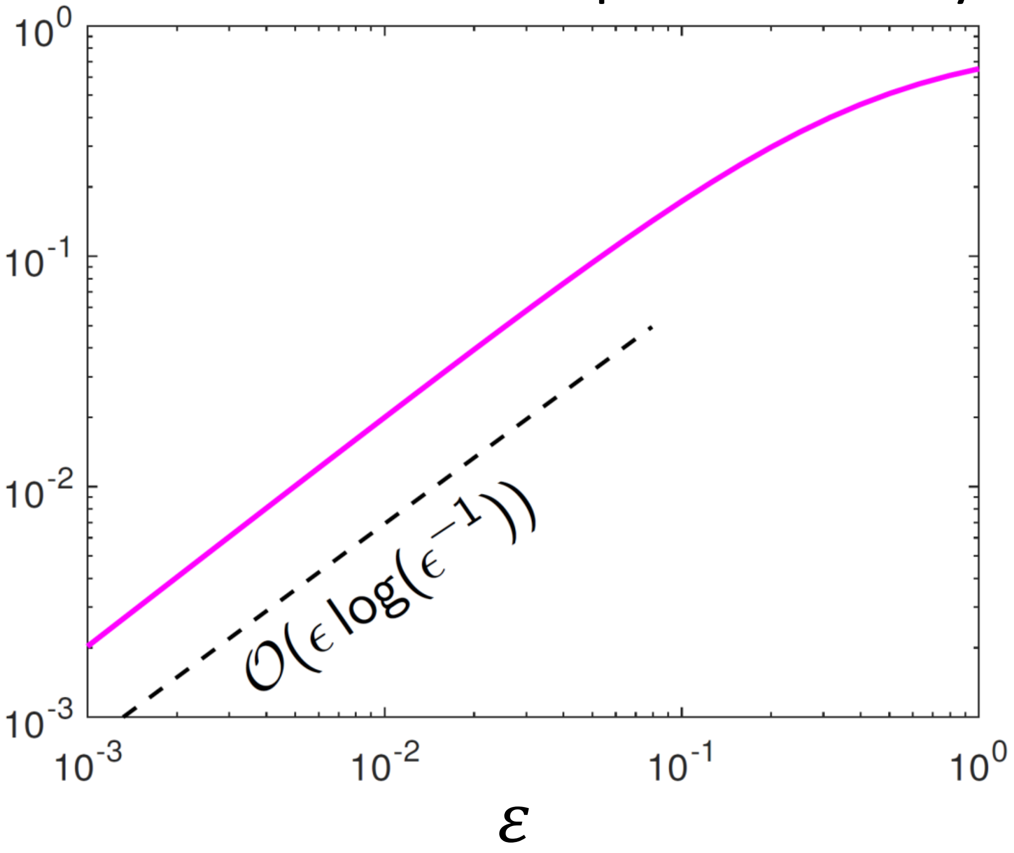




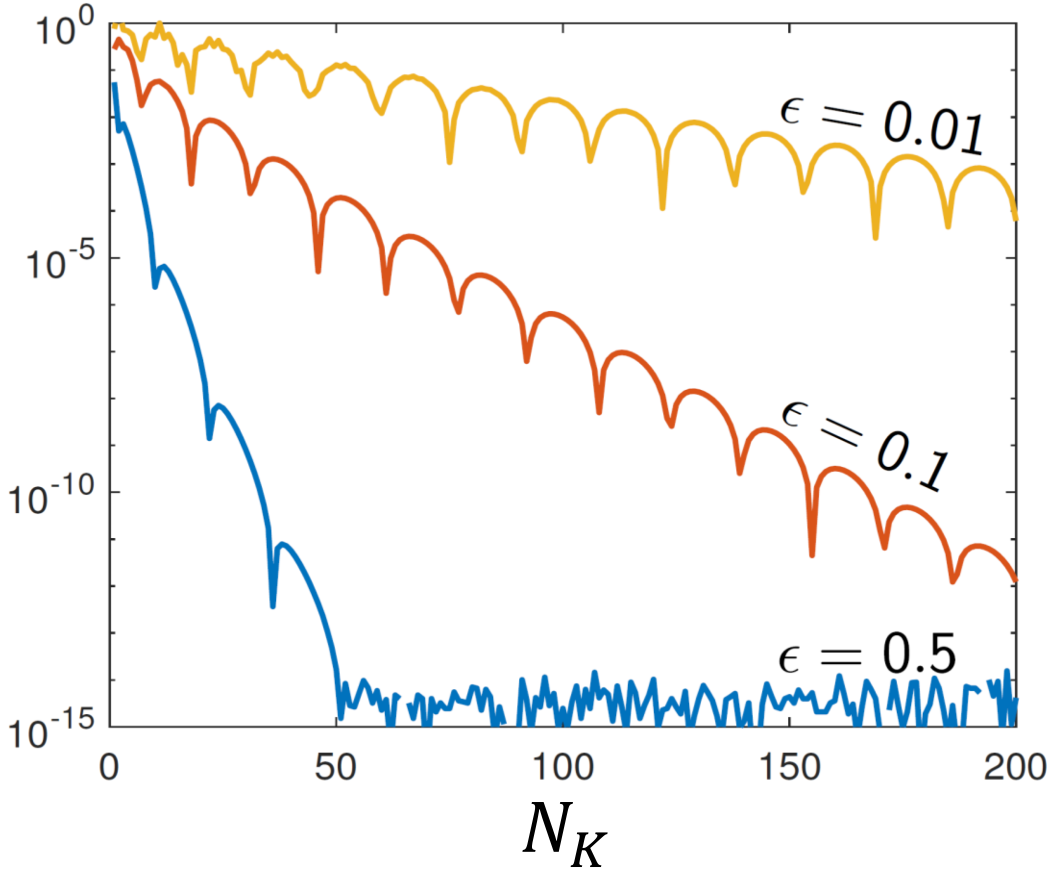
But … slow convergence

Problem: As $\varepsilon \downarrow 0$, error is $O(\varepsilon \log(1/\varepsilon))$ and $N_K(\varepsilon) \to \infty$.

Pointwise error for spectral density



Error due to discretization



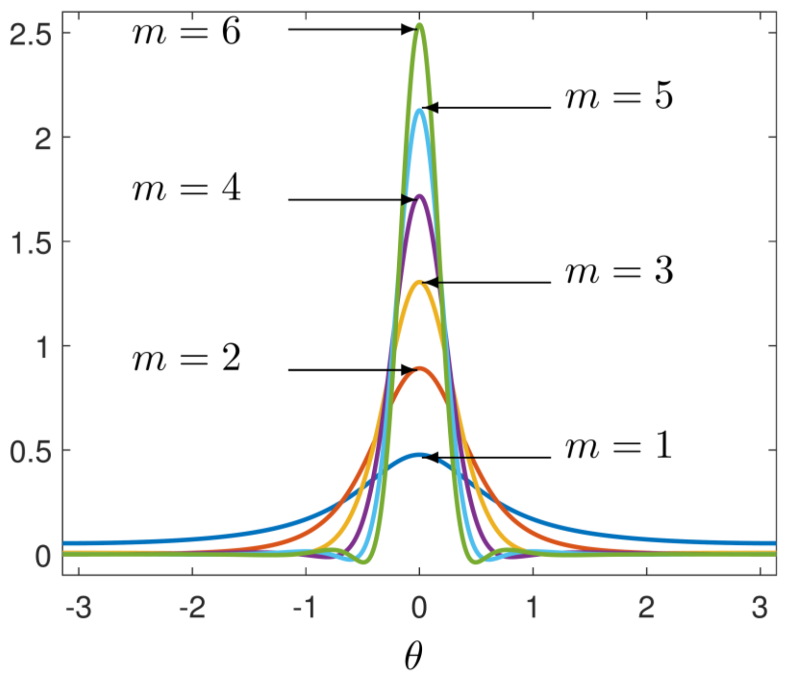
Small N_K critical in data-driven computations. Can we improve convergence rate?

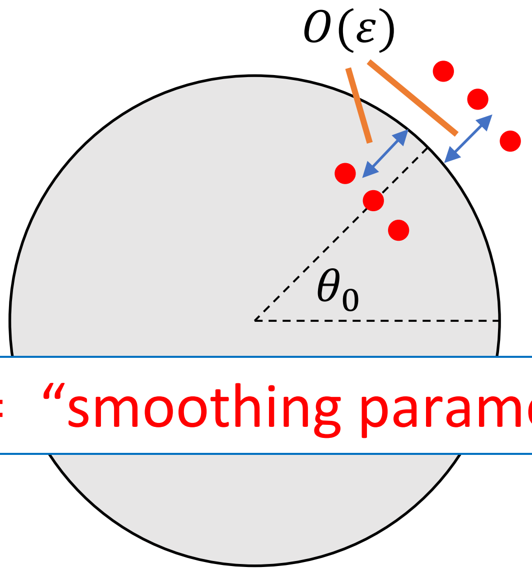
High-order rational kernels

mth order rational kernels:

$$K_\varepsilon(\theta) = \frac{e^{-i\theta}}{2\pi}\sum_{j=1}^{m}\left[\frac{c_j}{e^{-i\theta}-(1+\varepsilon\overline{z_j})^{-1}} - \frac{d_j}{e^{-i\theta}-\left(1+\varepsilon z_j\right)}\right]$$

Kernels



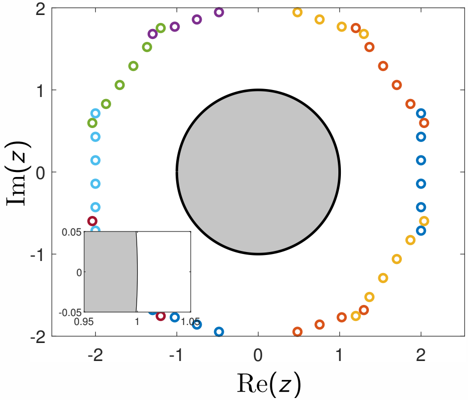


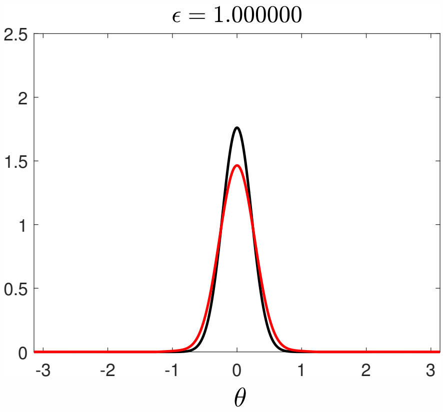
$\varepsilon =$ "smoothing parameter"

ResDMD computes with error control

$$[K_\varepsilon * \nu_g](\theta_0) = \sum_{j=1}^{m}\left[c_j\mathcal{C}_g\left(e^{i\theta_0}(1+\varepsilon\overline{z_j})^{-1}\right) - d_j\mathcal{C}_g\left(e^{i\theta_0}\left(1+\varepsilon z_j\right)\right)\right]$$

Smaller N_K (larger ε)





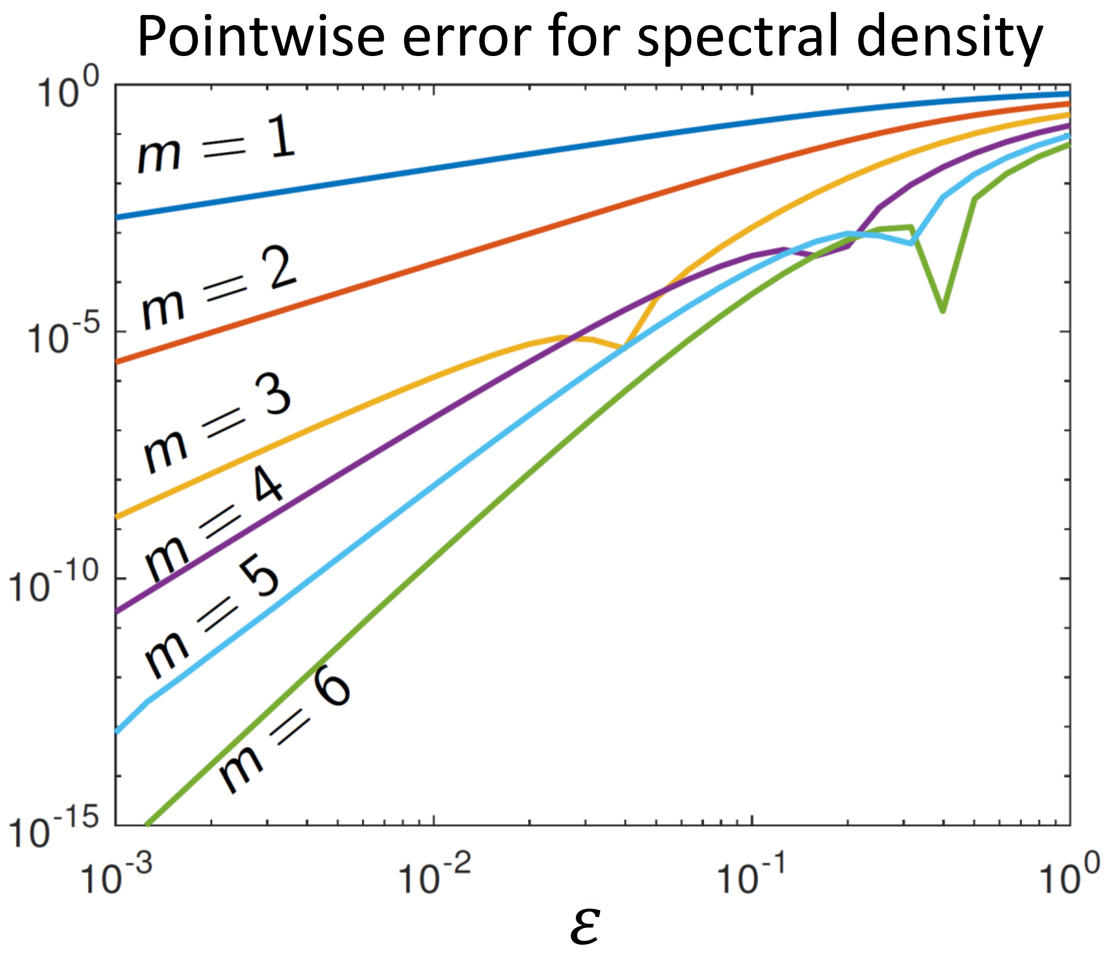
Convergence

Theorem: *Automatic* selection of $N_K(\varepsilon)$ with $O(\varepsilon^m \log(1/\varepsilon))$ convergence:

- Density of continuous spectrum ρ_g. (pointwise and L^p)
- Integration against test functions. (weak convergence)

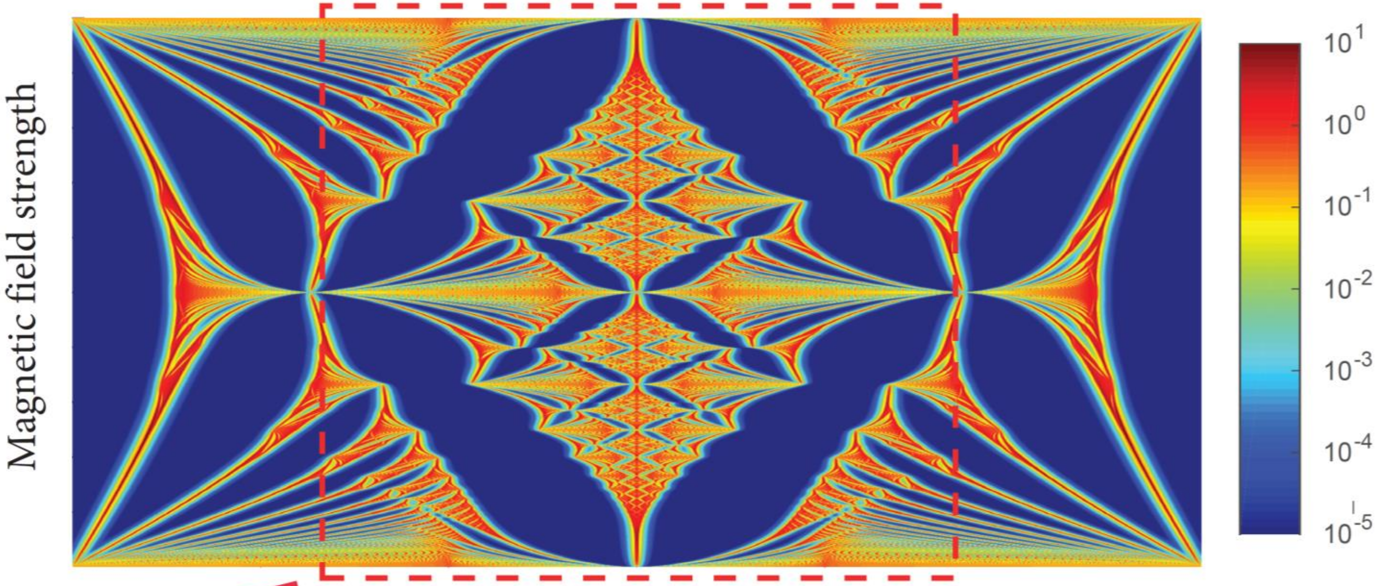
$$\int_{[-\pi,\pi]_{\text{per}}} h(\theta)\left[K_\varepsilon * \nu_g\right](\theta)\ \mathrm{d}\theta = \int_{[-\pi,\pi]_{\text{per}}} h(\theta)\,\mathrm{d}\nu_g(\theta) + O(\varepsilon^m \log(1/\varepsilon))$$

Also recover discrete spectrum.



- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.

Spectral measures of self-adjoint operators



Horizontal slice = spectral measure at constant magnetic field strength.

Software package

`SpecSolve` available at https://github.com/SpecSolve
Capabilities: ODEs, PDEs, integral operators, discrete operators.

- C., Horning, Townsend "*Computing spectral measures of self-adjoint operators,*" **SIAM Rev.**, 2021.

The Challenges

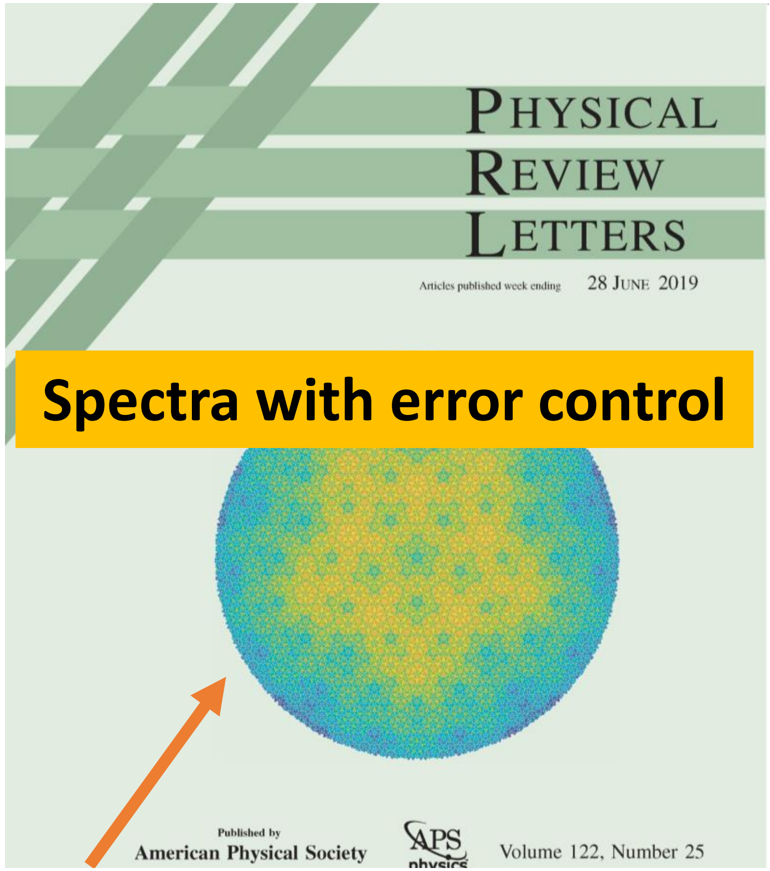
~~**1) "Too much":** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$~~ ✓

~~**2) "Too little":** Miss parts of $\mathrm{Spec}(\mathcal{K})$~~ ✓

~~**3) Continuous spectra.**~~ ✓

Verification: Is it right?

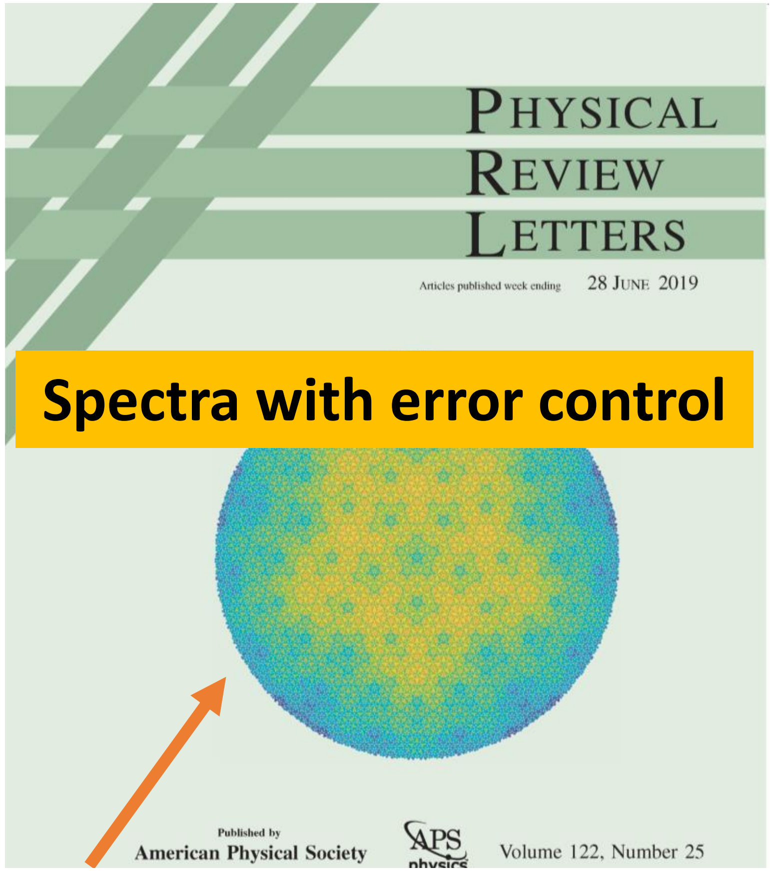
Is it right?
The importance of verification



E.g., ground state of quasicrystal

- C., Roman, Hansen, "*How to compute spectra with error control,*" **Phys. Rev. Lett.**, 2019.

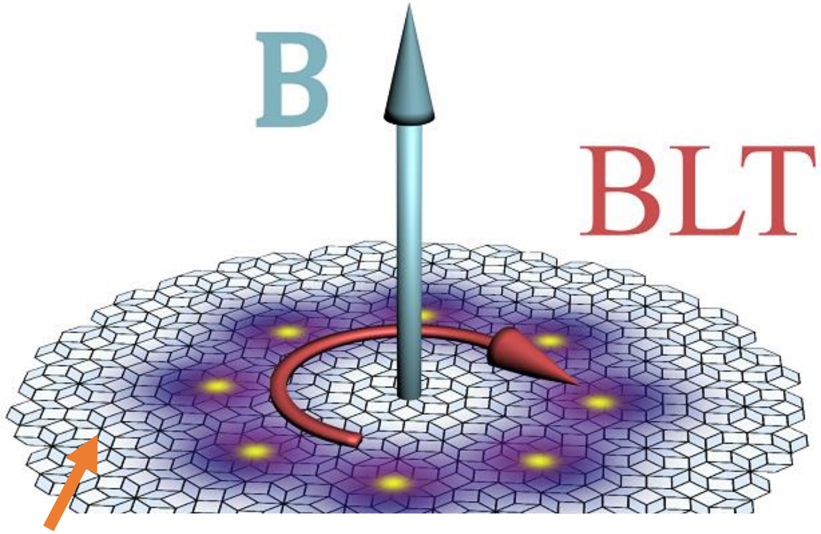
Is it right?
The importance of verification



Spectra with error control

E.g., ground state of quasicrystal

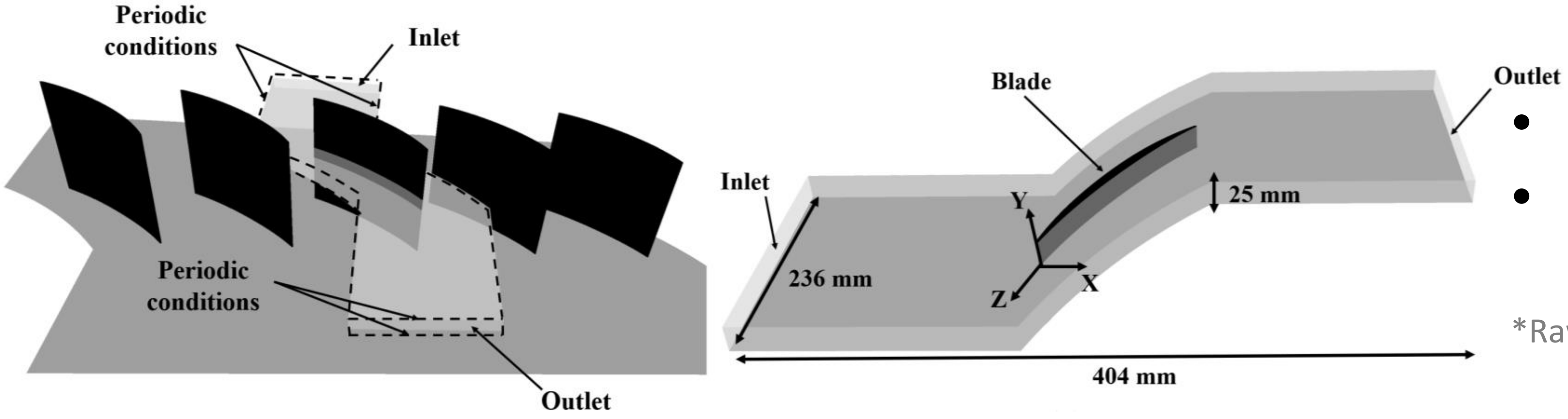
Certainty in computed spectral properties



E.g., new physical phenomena: **bulk localised transport** states

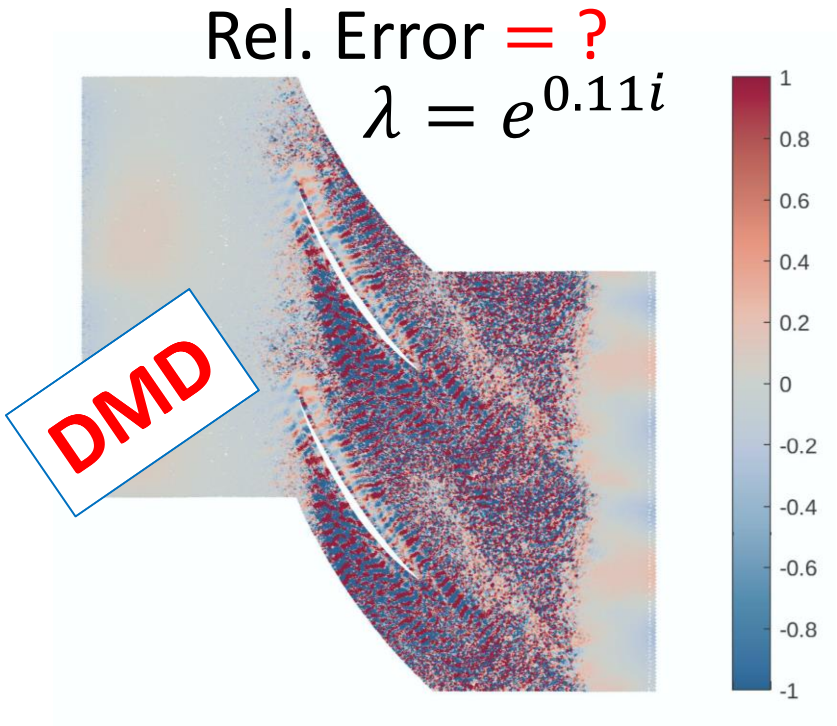
- C., Roman, Hansen, "*How to compute spectra with error control,*" **Phys. Rev. Lett.**, 2019.
- Johnstone, C., Nielsen, Öhberg, Duncan, "*Bulk Localised Transport States in Infinite and Finite Quasicrystals via Magnetic Aperiodicity,*" **Phys. Rev. B**, 2022.

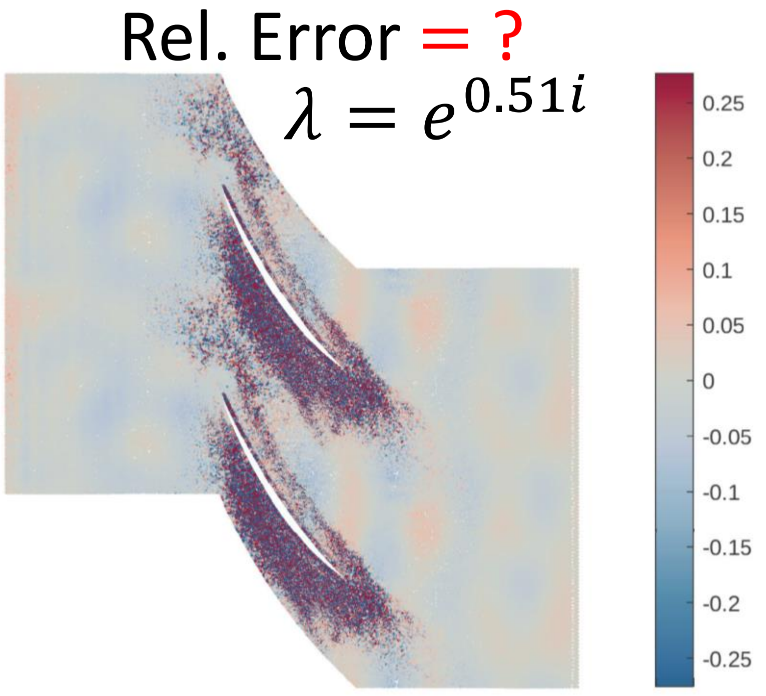
Example: Trustworthy computation for large d

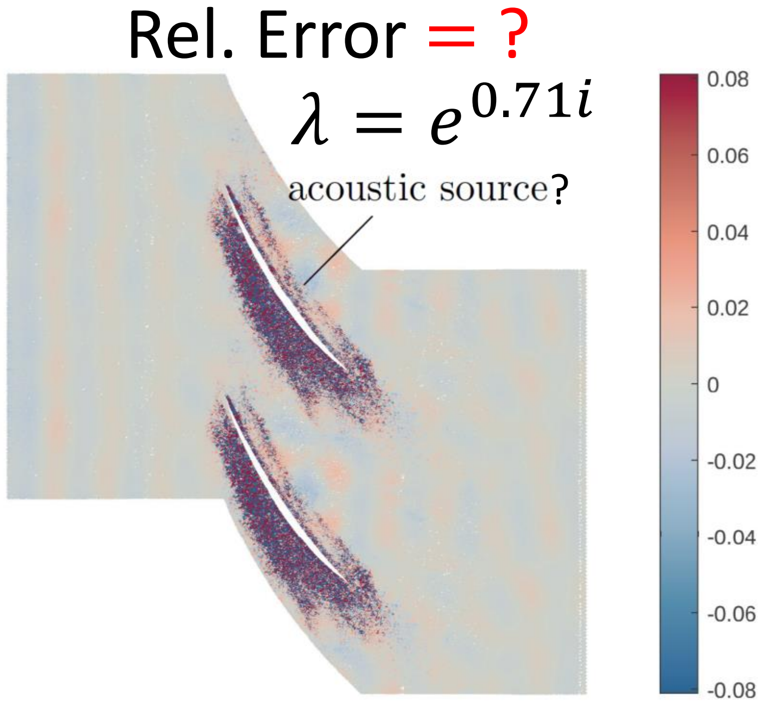


- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension (d) $\approx 300{,}000$ (number of measurement points)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

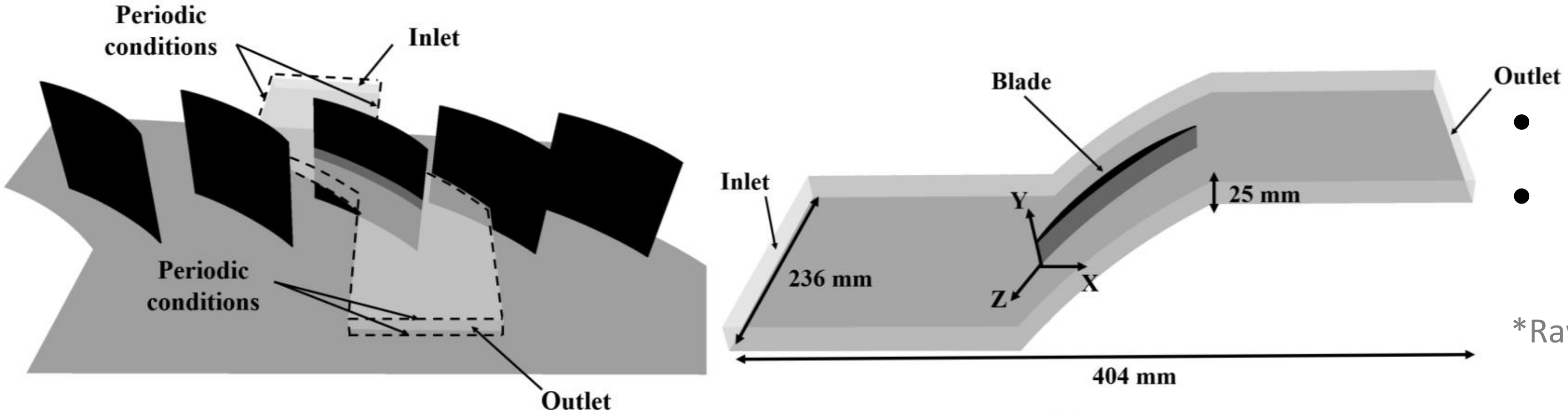






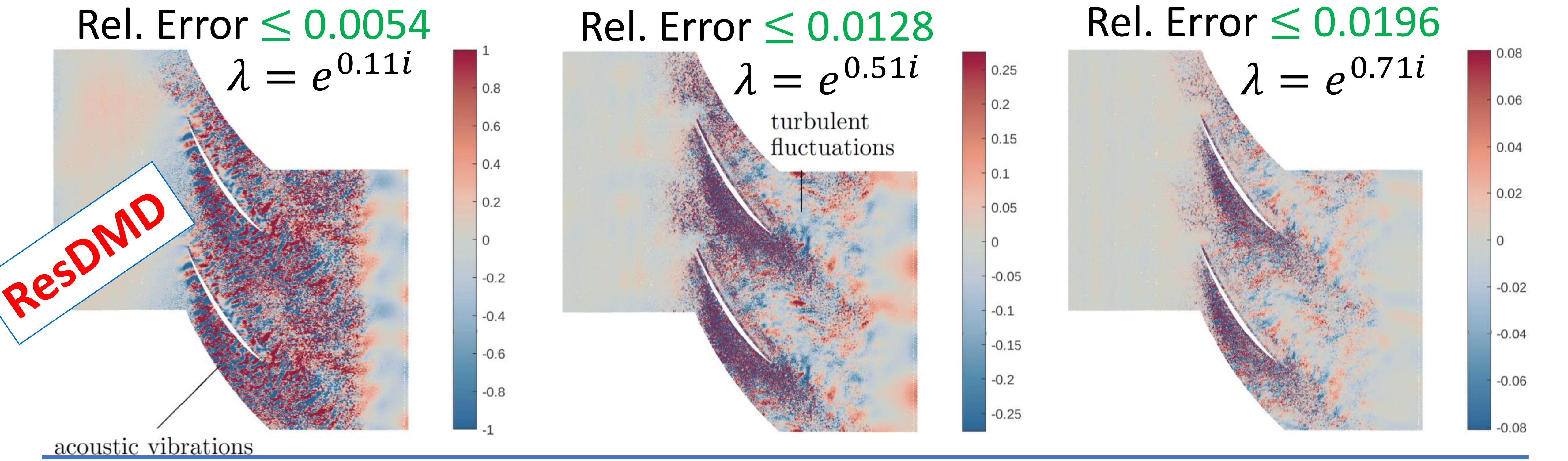
- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.

Example: Trustworthy computation for large d



- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension (d) $\approx 300{,}000$ (number of measurement points)

*Raw measurements provided by Stephane Moreau (Sherbrooke)



- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.

Large d ($\Omega \subseteq \mathbb{R}^d$): *robust* and *scalable*

Popular to learn dictionary $\{\psi_1, \ldots, \psi_{N_K}\}$

E.g., DMD with truncated SVD (linear dictionary, most popular), kernel methods (this talk), neural networks, etc.

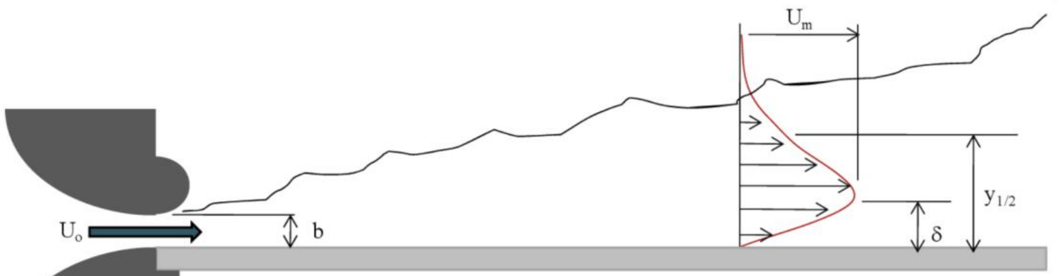
Q: Is discretisation span$\{\psi_1, \ldots, \psi_{N_K}\}$ large/rich enough?

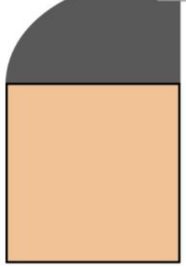
Above algorithms:

- Pseudospectra: $\{z_k : \tau_k < \varepsilon\} \subseteq \mathrm{Spec}_\varepsilon(\mathcal{K})$ **error control**
- Spectral measures: $\mathcal{C}_g(z)$ and smoothed measures **adaptive check**

$\Longrightarrow$ Rigorously ***verify*** learnt dictionary $\{\psi_1, \ldots, \psi_{N_K}\}$

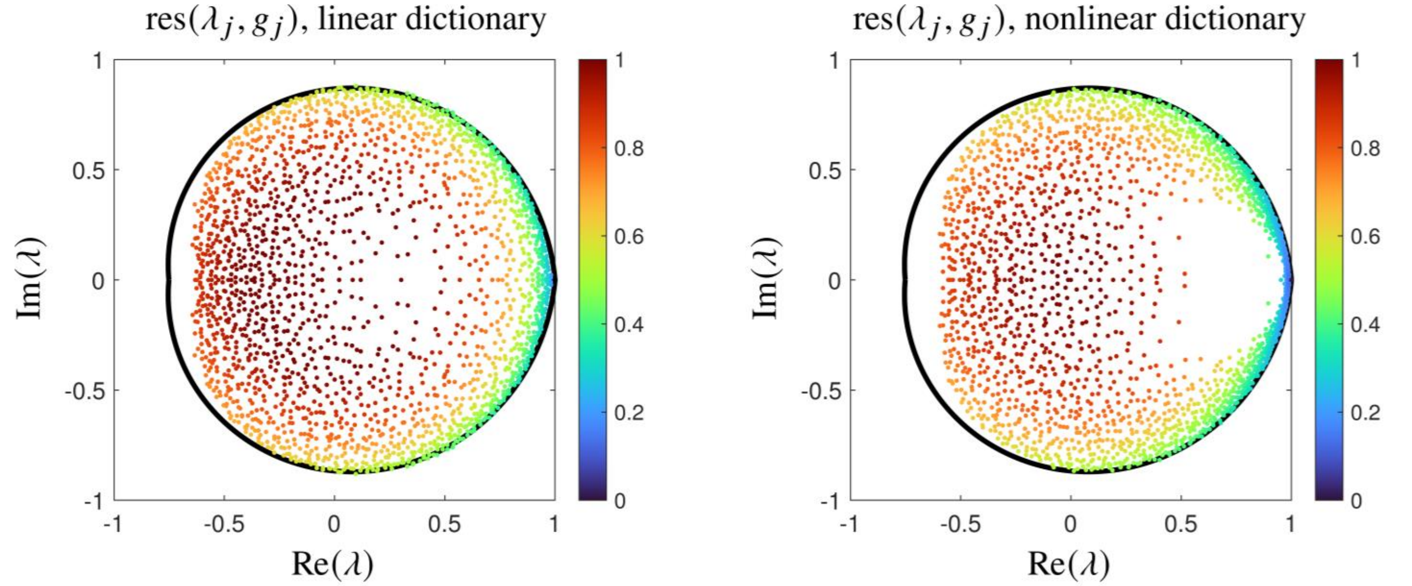
Example: Verify the dictionary



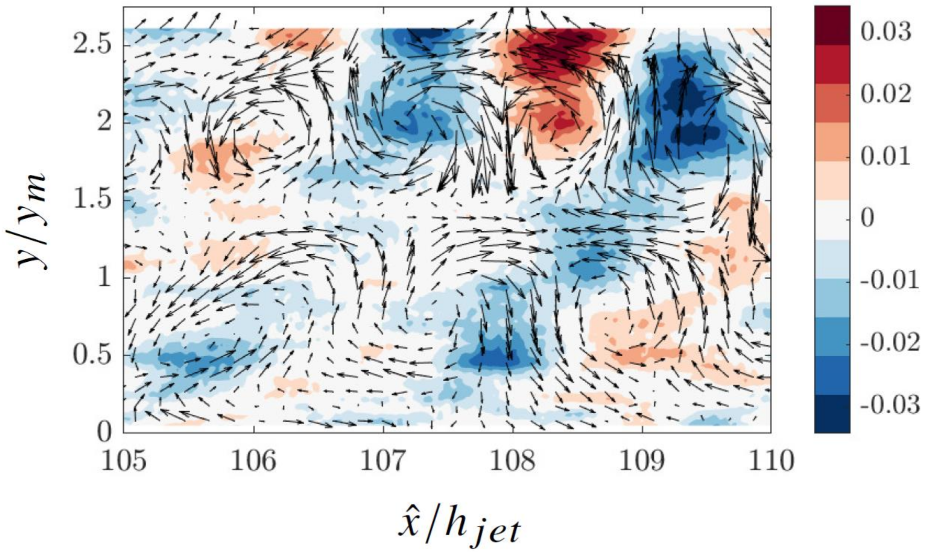


- Reynolds number $\approx 6.4 \times 10^4$
- Ambient dimension $(d) \approx 100{,}000$ (velocity at measurement points)

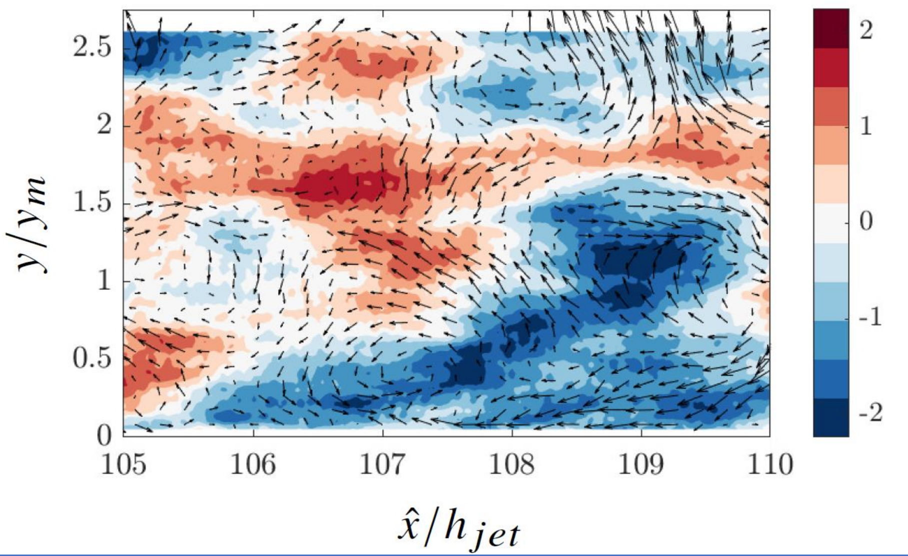
*Raw measurements provided by Máté Szőke (Virginia Tec



$\lambda = 0.9439 + 0.2458i$, error ≤ 0.0765

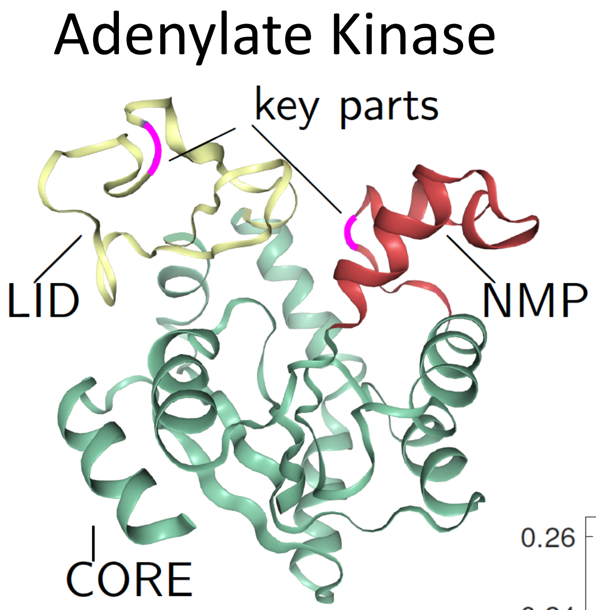


$\lambda = 0.8948 + 0.1065i$, error ≤ 0.1105



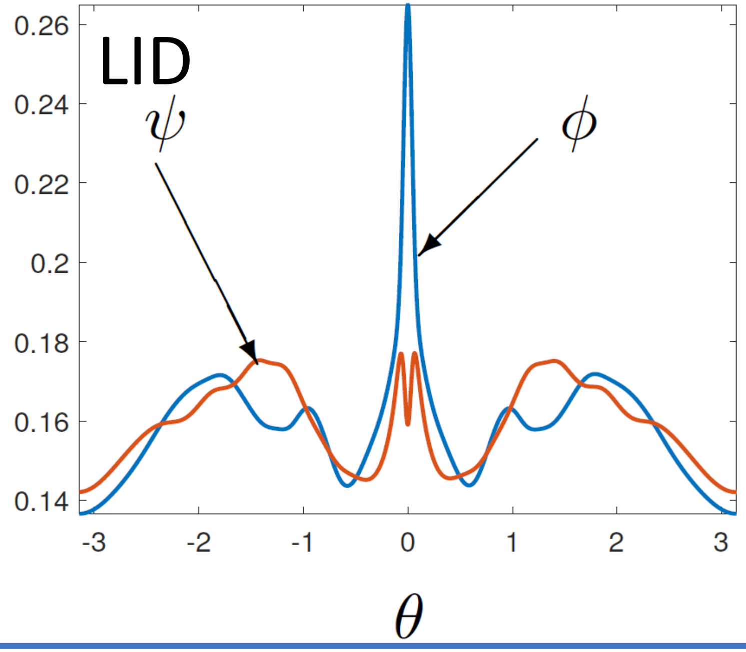
- C., Ayton, Szőke, "*Residual Dynamic Mode Decomposition,*" **J. Fluid Mech.**, under minor rev.

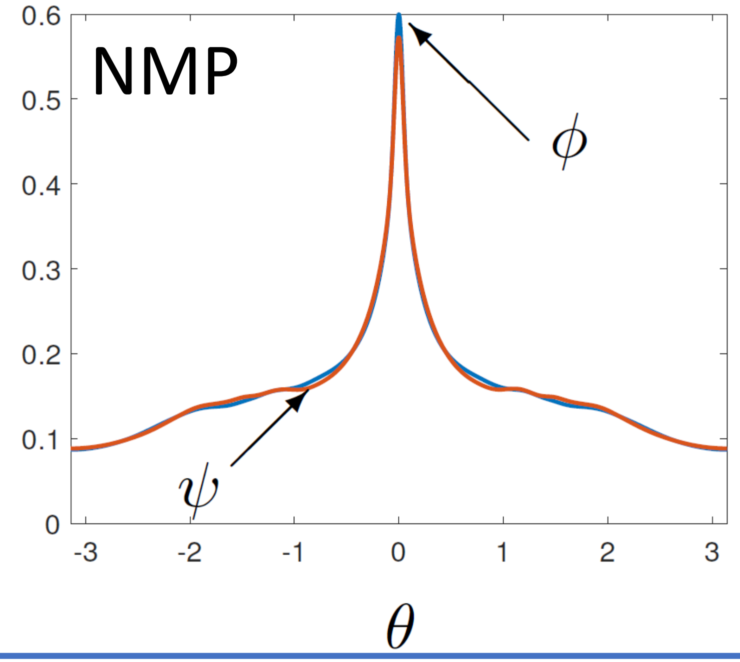
Example: Spectral measures in large d



- Ambient dimension (d) $\approx 20{,}000$ (positions and momenta of atoms)
- 6th order kernel (spec res 10^{-6})

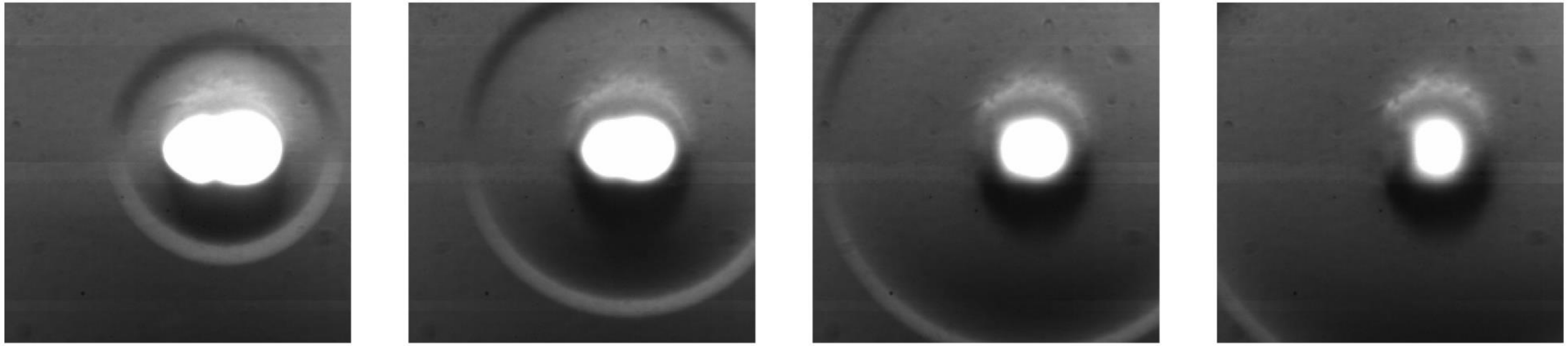
*Dataset: www.mdanalysis.org/MDAnalysisData/adk_equilibrium.html



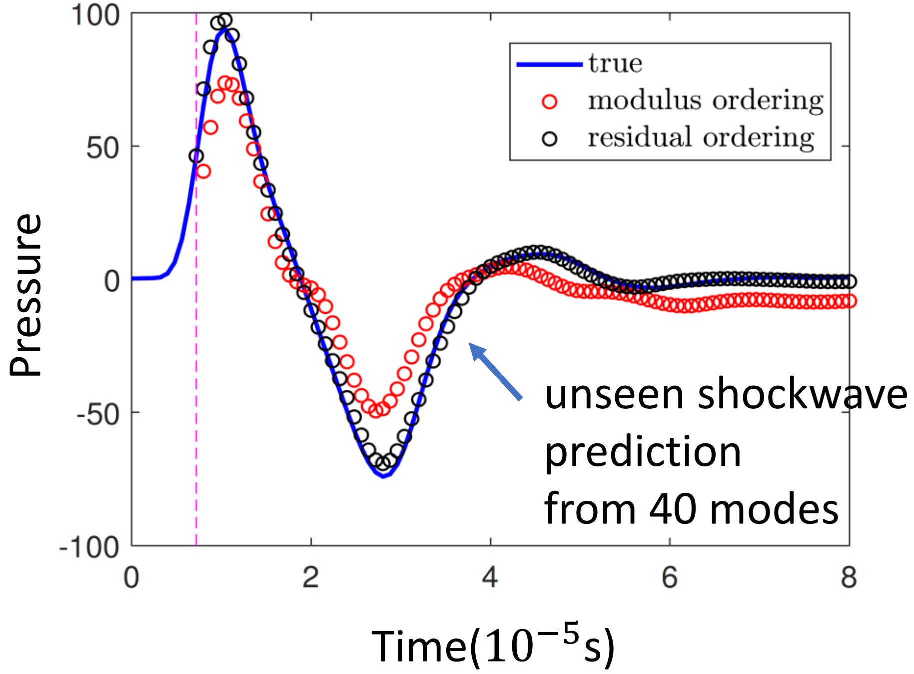


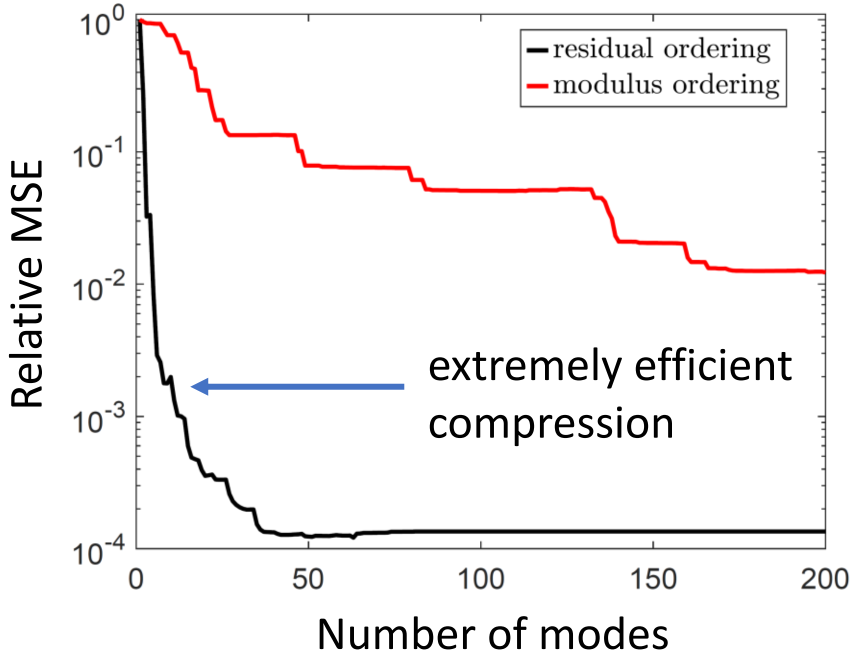
- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.

Example: Trustworthy Koopman mode decomposition



a) $t = 5\ \mu s$ b) $t = 10\ \mu s$ c) $t = 15\ \mu s$ d) $t = 20\ \mu s$





- C., Ayton, Szőke, "*Residual Dynamic Mode Decomposition,*" **J. Fluid Mech.,** under minor rev.

Wider programme

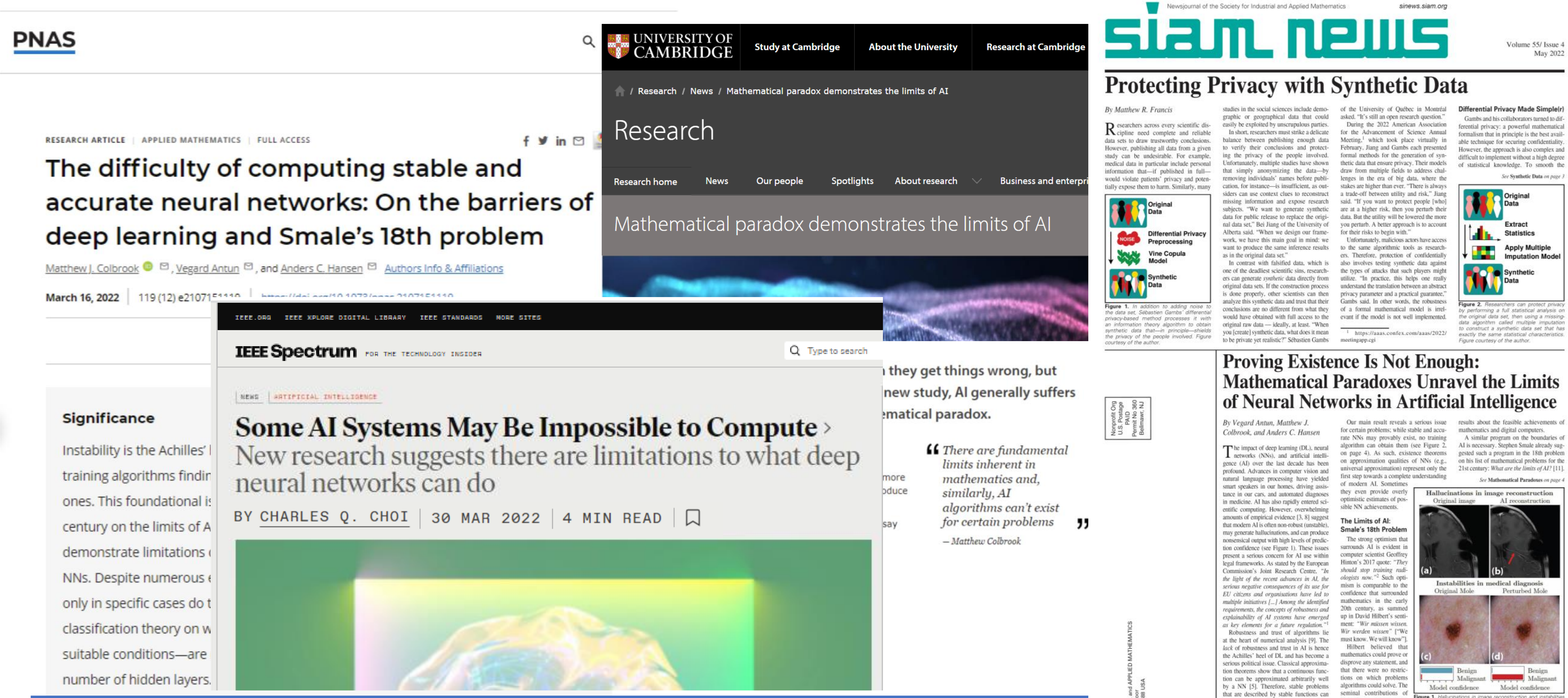
SCI provides needed assumptions

- <u>Infinite-dimensional computational analysis</u> $\Longrightarrow$ **Practical and rigorous algorithms.**
- <u>Solvability Complexity Index</u> $\Longrightarrow$ **Classify difficulty of problems, prove algorithms are optimal.**
- **Extends to:** Foundations of AI, optimization, computer-assisted proofs, and PDEs etc.

DATA SCIENCE + NUMERICAL ANALYSIS

- C., "*On the computation of geometric features of spectra of linear operators on Hilbert spaces*," **Found. Comput. Math.**, to appear.
- C., "*Computing spectral measures and spectral types*," **Comm. Math. Phys.**, 2021.
- C., Horning, Townsend "*Computing spectral measures of self-adjoint operators*," **SIAM Rev.**, 2021.
- C., Roman, Hansen, "*How to compute spectra with error control*," **Phys. Rev. Lett.**, 2019.
- C., Hansen, "*The foundations of spectral computations via the solvability complexity index hierarchy*," **J. Eur. Math. Soc.**, 2022.
- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem*," **Proc. Natl. Acad. Sci. USA**, 2022.
- C., "*Computing semigroups with error control*," **SIAM J. Numer. Anal.**, 2022.
- C., "*The mpEDMD Algorithm for Data-Driven Computations of Measure-Preserving Dynamical Systems*," preprint.
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, "*On the solvability complexity index hierarchy and towers of algorithms*," arXiv, 2020.
- Smale, "*The fundamental theorem of algebra and complexity theory*," **Bull. Amer. Math. Soc.**, 1981, *36 pp.*
- McMullen, "*Families of rational maps and iterative root-finding algorithms*," **Ann. of Math.**, 1987, *27 pp.*

Example: Barriers of deep learning



- C., Antun, Hansen, *"The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,"* **Proc. Natl. Acad. Sci. USA.**

Interested? Get in touch (e.g., I'll be around this afternoon)!

"One of great joys of doing mathematics is working with inspiring and brilliant people!"
- Arieh Iserles

Some future directions:

- ResDMD + control $\Longrightarrow$ error control?
- Embed & learn symmetries (e.g., check out the algorithm mpEDMD).
- Forecasting with error bounds.
- Koopmanism meets neural nets (and vice versa).
- Foundations results for dynamical systems (i.e., impossibility results)?
- Further barriers in deep learning.
- <u>Functional analysis</u> meets <u>data science</u> meets <u>numerical analysis</u>!

Opportunities to collaborate, visit Cambridge, grad students & beyond!

Summary: rigorous data-driven Koopmanism!

- "Too much" or "Too little"

  Idea: New matrix for residual ⇒ **ResDMD** for computing spectra.

- Continuous spectra and spectral measures:

  Idea: Convolution with rational kernels via resolvent and **ResDMD**.

- Is it right?

  Idea: Use **ResDMD** to verify computations. E.g., learned dictionaries.

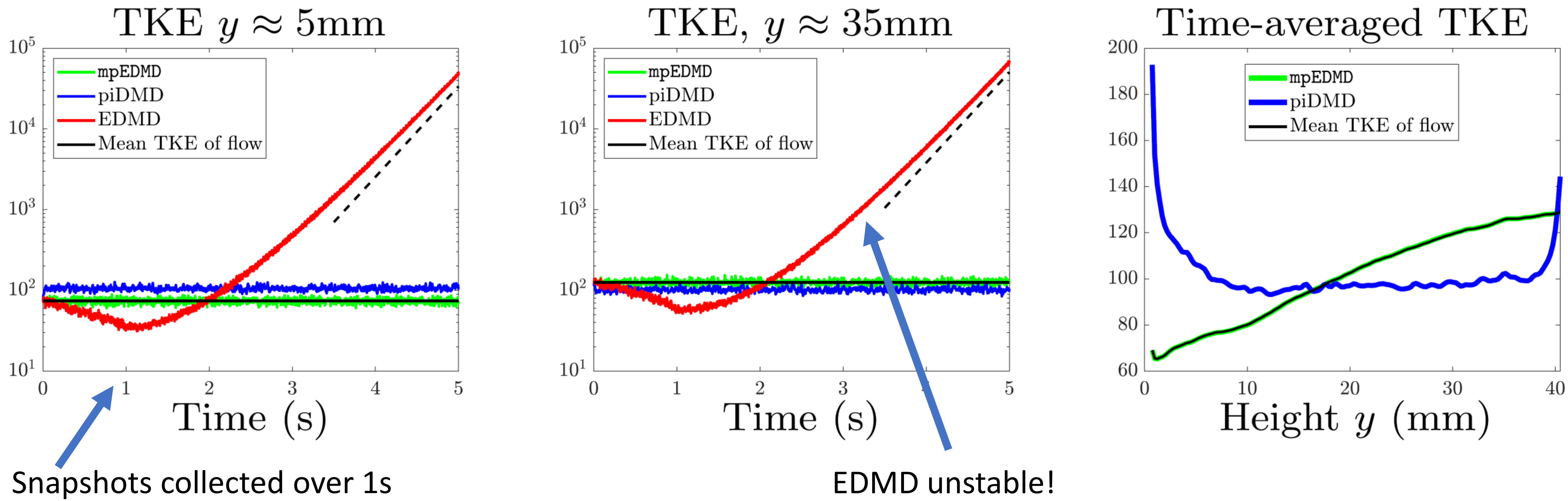
Code:

https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition

Additional slides…

measure-preserving EDMD…

- Polar decomposition of $\mathcal{K}$. Easy to combine with any DMD-type method!
- Converges for spectral measures, spectra, Koopman mode decomposition.
- Measure-preserving discretization for arbitrary measure-preserving systems.



- C., *"The mpEDMD Algorithm for Data-Driven Computations of Measure-Preserving Dynamical Systems,"* arXiv 2022.

Solvability Complexity Index Hierarchy

Class $\Omega \ni A$, want to compute $\Xi\colon \Omega \to (\mathcal{M}, d)$ metric space

- Δ_0: Problems solved in finite time (v. rare for cts problems).
- Δ_1: Problems solved in "one limit" with full error control:

$$d(\Gamma_n(A), \Xi(A)) \le 2^{-n}$$

- Δ_2: Problems solved in "one limit":

$$\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$$

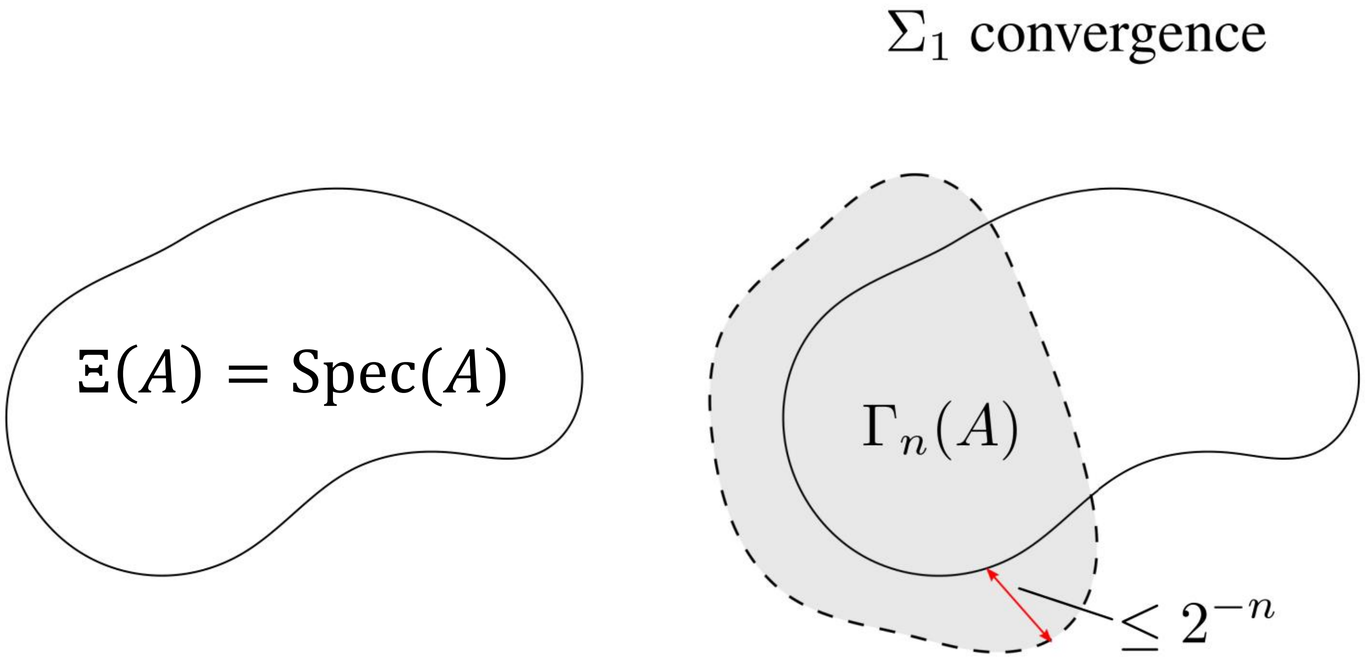
- Δ_3: Problems solved in "two successive limits":

$$\lim_{n\to\infty} \lim_{m\to\infty} \Gamma_{n,m}(A) = \Xi(A)$$

$\vdots$

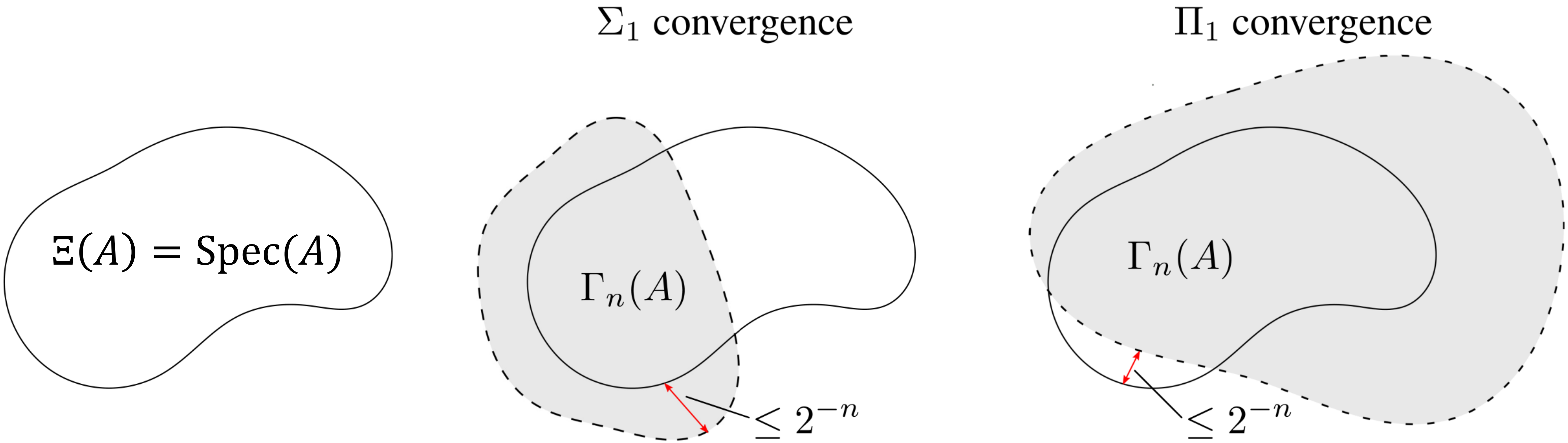
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, "*On the solvability complexity index hierarchy and towers of algorithms,*" preprint.
- Hansen, "*On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators,*" **J. Amer. Math. Soc.**, 2011.
- McMullen, "*Families of rational maps and iterative root-finding algorithms,*" **Ann. of Math.**, 1987.
- Doyle, McMullen, "*Solving the quintic by iteration,*" **Acta Math.**, 1989.
- Smale, "*The fundamental theorem of algebra and complexity theory,*" **Bull. Amer. Math. Soc.**, 1981.

Error control for spectral problems



- Σ_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$, $\max_{z\in\Gamma_n(A)} \mathrm{dist}\big(z, \Xi(A)\big) \leq 2^{-n}$

Error control for spectral problems



- Σ_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A),\ \max_{z\in\Gamma_n(A)} \mathrm{dist}\big(z, \Xi(A)\big) \leq 2^{-n}$
- Π_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A),\ \max_{z\in\Xi(A)} \mathrm{dist}(z, \Gamma_n(A)) \leq 2^{-n}$

Such problems can be used in a proof!

Small sample of classification theorems

Increasing difficulty

Error control

$$\begin{array}{ccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 \\
\| & & & \subsetneq & & \subsetneq & & \subsetneq & & \subsetneq & & \subsetneq & & \cdot^{\cdot^{\cdot}} \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 \ldots \\
\| & & & \subsetneq & & \subsetneq & & \subsetneq & & \subsetneq & & \subsetneq & & \cdot_{\cdot_{\cdot}} \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3
\end{array}$$

Small sample of classification theorems

Increasing difficulty

Error control

$$
\begin{tikzcd}[column sep=small]
\Pi_0 & & \Pi_1 & & \Pi_2 & & \Pi_3 & \iddots \\
\Delta_0 \arrow[u, equal] \arrow[d, equal] \arrow[r, phantom, "\subsetneq"] & \Delta_1 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_1 \cup \Pi_1 \arrow[r, phantom, "\subsetneq"] & \Delta_2 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_2 \cup \Pi_2 \arrow[r, phantom, "\subsetneq"] & \Delta_3 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_3 \cup \Pi_3 & \ldots \\
\Sigma_0 & & \Sigma_1 & & \Sigma_2 & & \Sigma_3 & \ddots
\end{tikzcd}
$$

Spectra of $\mathcal{K}$

Continuous spectra of $\mathcal{K}$ (different regularity assumptions)

Small sample of classification theorems

Increasing difficulty

Error control

Spectra of compact operators

$$\begin{array}{ccccccccccccccc}
\Pi_0 & & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 & \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \ddots \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 & \ldots \\
\| & & & \rotatebox{45}{$\subsetneq$} & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \iddots \\
\Sigma_0 & & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3 &
\end{array}$$

Spectra of $\mathcal{K}$

Continuous spectra of $\mathcal{K}$ (different regularity assumptions)

Small sample of classification theorems

Increasing difficulty

Error control

Spectra of compact operators

Spectra of Schrödinger*
(different potential classes)

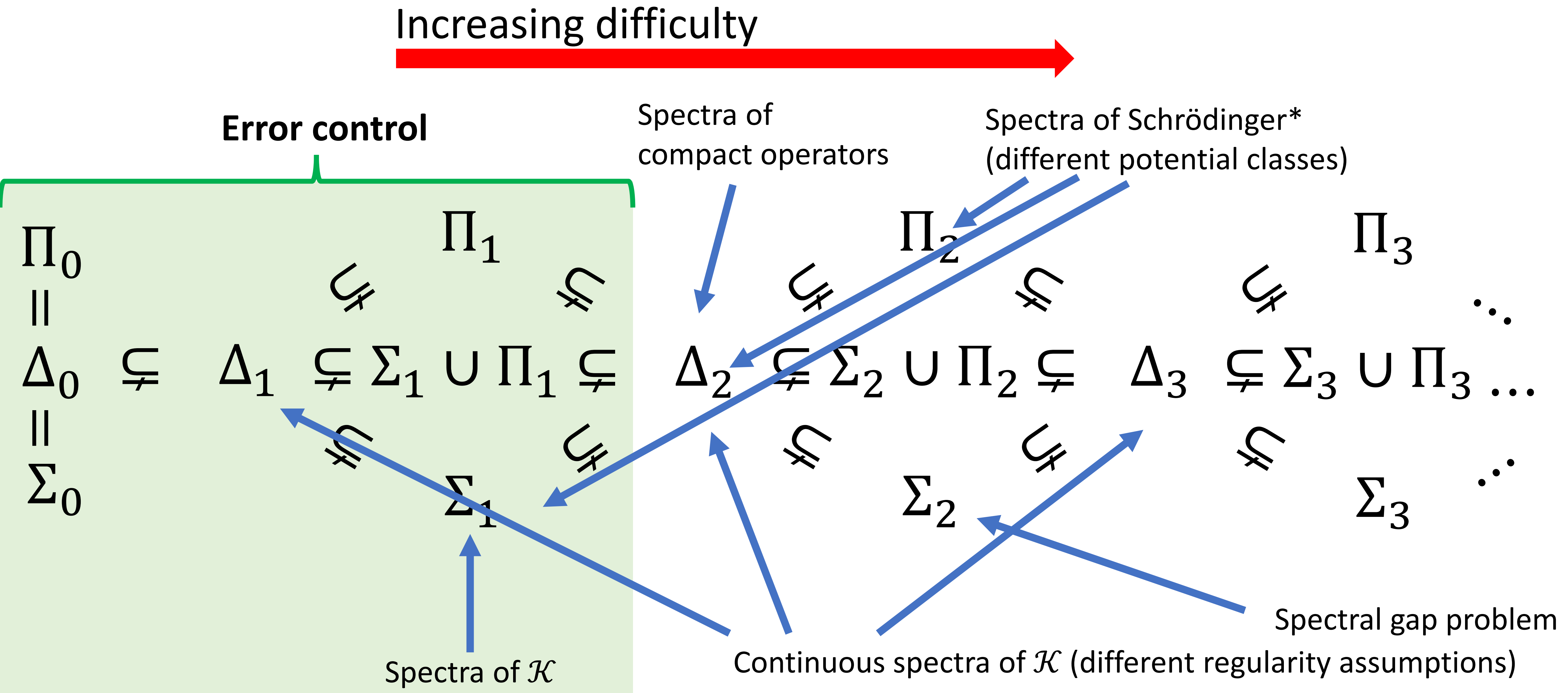
$$
\begin{array}{ccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 \\
\| & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \ddots \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 \ \dots \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \iddots \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3
\end{array}
$$

Spectra of $\mathcal{K}$

Continuous spectra of $\mathcal{K}$ (different regularity assumptions)

*_Open problem of Schwinger_: “The special canonical group,” “Unitary operator bases,” PNAS, 1960.

Small sample of classification theorems



Open problem of Schwinger: "The special canonical group," "Unitary operator bases," PNAS, 1960.

Small sample of classification theorems

Increasing difficulty

Error control

Spectra of compact operators

Spectra of Schrödinger* (different potential classes)

Spectral stability

$$
\begin{array}{ccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 \\
\| & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \ddots \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 \ \dots \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \iddots \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3
\end{array}
$$

Spectra of $\mathcal{K}$

Spectral gap problem

Continuous spectra of $\mathcal{K}$ (different regularity assumptions)

**Open problem of Schwinger*: “The special canonical group,” “Unitary operator bases,” PNAS, 1960.